Dear Editor,

we thank you very much for editing our manuscript. We highly valued the reviewer comments, and we believe that we could strongly improve the manuscript based on them. We provide the updated manuscript along with the track-changes. The detailed point-to-point replies can be found below.

The largest change was the restructuring of the manuscript, clearer distinguishing between the submodel description, the methods and the results. Unfortunately, this leads to a slightly confusing track-changes file. The major according structural changes are provided in the replies to the general comments below. All further major changes are described below, in addition to the track-changes file.

Note, that we decided to show the sensitivity on the lower integration limit, and the information on the optimization for the degassing volcanic event in the appendix instead of the supplement (as we had stated in the replies).

Thank you again and best regards,

Matthias Kohl (on behalf of all co-authors)

Reply to comments of RC1:

We thank the reviewer for the time to review our manuscript, and especially the very detailed and constructive feedback. After the initial reply on the general comments, we now want to address all comments in detail. Here we also repeat and extend the answers to the general comments already posted in the open discussion for completeness, now also considering the comments of the second reviewer and the revised manuscript. We report the comments (grey, bold) along with our replies (blue).

The submitted manuscript contains the description of a new submodule within MESSy. The submodule is described in detail in section 2. Section 3 provides a description of the model setup for the validation of the submodel. For model validation the 2011 Nabro and 2018 Kilauea SO_2 emission plumes were evaluated with different sensitivity simulations and with a fine-resolution simulation, respectively. In chapter 4 the explosive volcanic SO_2 injections of the time period 2007 to 2011 was simulated based on a 3-D observational emission dataset which was improved based on the findings of chapter 3.

The submitted manuscript describes a new submodule for emission of volcanic substances and provides some validation for the evolution of SO_2 plumes after volcanic eruptions. However, in my point of view, the manuscript has some substantial flaws, which I explain below and which need to be addressed. Even though topic wise the manuscript would be a good fit for publication in GMD, I do not think that quality wise this manuscript deserves publication in this journal.

We thank the reviewer for the assessment of our manuscript. We agree that there are some shortcomings that need to be addressed, however, we disagree on the overall evaluation. We will discuss the comments in detail below.

The main environmental impact of large explosive volcanic eruption on a global scale (heterogeneous chemistry, cooling of the surface, local warming of the stratosphere) are effects from sulfuric acid aerosols and not from SO_2. The manuscript almost entirely focuses on the simulation of SO_2 plumes and thus, a crucial aspect of volcanic emission plumes is missing. I know that for the accurate simulation of aerosol burden and their effects it is important to accurately simulate the SO2 plume in the first place, however, if this was the goal of this study this needs to be motivated and explained in the introduction.

We agree with the reviewer that the main environmental impact of volcanic eruptions comes from sulfur aerosol (we prefer to talk about "sulfur aerosol" in contrast to "sulfuric acid aerosol", as it also includes sulfate and bi-sulfate species such as ammonium (bi-)sulfate; see also later specific comment). However, as the reviewer pointed out, these sulfur aerosols are a result of volcanic SO_2 emissions. We also stated this twice in the introduction of the submitted manuscript: "Moreover, SO_2 undergoes oxidation to form sulfuric acid (H_2SO_4), which rapidly converts to the aerosol phase forming sulfate under most atmospheric conditions." and "When volcanic SO_2 emissions reach the stratosphere, the subsequently formed sulfate aerosols enhance the stratospheric aerosol burden and are distributed widely across the globe". Therefore, for the evaluation of the EVER submodel (handling the initial volcanic emissions), SO_2 is the correct tracer that needs to be investigated and evaluated. Nevertheless, the effects of sulfur aerosol are described in the introduction, as they have the most important impact, and they

result from volcanic SO_2 emissions.

Indeed, the goal of the study is to present a new submodel that handles emissions of gaseous and aerosol species, and the accurate simulation of volcanic SO_2. The subsequent formation of sulfuric acid and aerosols is a consequence of chemistry, thermodynamics and microphysics. This is handled by the other submodels present in the MESSy model system (e.g. MECCA (Sander et al., 2019) and GMXe (Pringle et al., 2010)), which have been evaluated elsewhere and therefore are not subject of this paper. We also state this in the introduction: "Our primary aim is to provide recommendations regarding the implementation of stratospheric SO_2 injections from volcanic eruptions within the MESSy framework ...", "We evaluate the EVER submodel using SO_2 emissions from the 2011 Nabro explosive volcanic eruption ..." and "Based on the stratospheric volcanic SO_2 emission inventory developed by Schallock et al. (2023), we establish a default setup for the EVER submodel ...". In the revised version, we make the main goal of the study clearer in the introduction, and distinguish it from previously published studies.

Regarding the comment "a crucial aspect of volcanic emission plumes is missing": We disagree on that. The EVER submodel is only responsible for the emission of the initial plume species. Sulfur aerosols are not part of the "volcanic emission plume". They form in the evolution of the plume. Standard aerosol simulations with EMAC (or MESSy in general) already included the chemistry, thermodynamics and microphysics leading to the formation of sulfur aerosol from SO_2 (see for example Brühl et al. (2018) and Schallock et al. (2023)), however it was missing a standard setup for including the initial SO_2 emissions, which we wanted to address.

As we agree, that the mid-term SO_2 mixing ratios (Sect. 3.3.2 of the submitted manuscript) and the resulting aerosol evaluated later depend on the SO_2 removal and microphysics, we talk about their handling in more detail in the revised version (see also comments below), but this is a more technical comment in our opinion. Moreover, the chemical mechanism, including H_2SO_4 formation, is already provided as supplement, which we refer to in the manuscript. We also want to add, that the investigation of the resulting AOD is just an additional validation, and should not be seen as a volcanic aerosol study. We apologize, if we implied otherwise, and we clarified this in the revised version.

To me many important aspects of a paper are not sufficiently explained: What is the novelty of this study? What can the model do what other models can't do? Why is it important to simulate the emissions of SO_2 accurately? Why do we need yet another model which is capable of simulating volcanic injections? What have other models shown and how does your model compare to other models?

This manuscript is a documentation of a new submodel, along with its evaluation and the development of a default setup. It is not a study, aiming to provide new results on volcanic eruptions. This is well within the scope of GMD papers. From the "Aims and Scope" of GMD papers (`https://www.geoscientific-model-development.net/about/aims_and_scope.html`) amongst others:

- "geoscientific model descriptions, from statistical models to box models to GCMs"

- "development and technical papers, describing developments such as new parameterizations or technical aspects of running models such as the reproducibility of results"

We chose the manuscript type "Development and technical papers" (from `https://www.geoscientific-`

model-development.net/about/manuscript_types.html): "These papers describe technical developments relating to model improvements such as the speed or accuracy of numerical integration schemes as well as new parameterisations for processes represented in modules. [...] Development and technical papers usually include a significant amount of evaluation against standard benchmarks, observations, and/or other model output as appropriate."

We believe that most of the raised questions are already sufficiently answered within the scope of a GMD paper.

- **"What is the novelty of this study?"** In MESSy, we so far only had a more or less generic submodel (TREXP, Jöckel et al. (2010)) for point or column like emissions in the atmosphere. This submodel, however, was made for different purposes, and missing features which are particularly required for emissions from volcanic eruptions (e.g. the possibility to specify vertical emission profile shapes or primary aerosol emissions as needed for volcanic ash). To avoid increasing the complexity of TREXP, we decided to implement EVER. We describe this in lines 125ff. In previous studies a lot of manual work was necessary to consider explosive volcanoes and other injections into the stratosphere, and there was no standard setup covering stratospheric volcanic eruptions. In addition, we provide recommendations on the implementation of the initial SO_2 emissions of explosive volcanic eruptions, based on the available emission inventories, as Brodowsky et al. (2021) and Quaglia et al. (2023) found, that the different emission databases lead to vastly differing results. This can be mostly attributed to the unclear fraction of the volcanic plume in the stratosphere. We now provide more information on that in the abstract, introduction, discussion and conclusions in the revised manuscript.

- **"What can the model do what other models can't do?"** Indeed, other models can also handle volcanic eruptions; however, in the MESSy model system, it was not possible so far. So, we don't see any contradiction here. We present our new submodel and evaluate it using observations. The implementation differs from other models being specific to the MESSy interface. Additionally, we studied the correct emission location, timing and stratospheric mass that has to be implemented, while previous studies (e. g. Brodowsky et al., 2021) only used some assumptions, on which mass fraction from the emission inventory of Brühl et al. (2018) has to be emitted in the stratosphere, leading to vastly differing stratospheric SO_2 burdens. We refer to other model developments in more detail in the introduction and the discussion of the revised version.

- **"Why is it important to simulate the emissions of SO_2 accurately?"** Details about the importance of sulfur aerosol resulting from volcanic SO_2 emissions are listed in the introduction. This implies the importance of accurately simulating SO_2 emissions. The results also show, how differing SO_2 emission parameters lead to differing results. We include these findings in the abstract and in the conclusion in the revised version.

- **"Why do we need yet another model which is capable of simulating volcanic injections?"** As there was none in MESSy, and other available models were not sufficient for our needs, this paper described a further development of the code. We made it clearer in the introduction, why we chose an own implementation, and thank the reviewer for this suggestion. The most important reason is the handling of the position, timing and stratospheric mass of the initial SO_2 emission, that was not investigated in previous volcanic timeline studies. We added more detail in the introduction, where we discuss

available volcanic models and emission inventories.

- **"What have other models shown and how does your model compare to other models?"** The main goal of the paper is to achieve reasonable agreement with observations to provide a standard setup. For that reason, we mostly focused on the observational side as the validation by observations is more important than comparisons to other model simulations in our opinion. However, we agree with the reviewer, that it is important to discuss other models, and we highlighted the differences and similarities to other model studies in the revised version in the introduction (see also next comment).

To avoid further confusion, we generally provide a clearer description of the nature and the goal of the study in the introduction, as mentioned in the points above.

I think the manuscript does not provide/discuss a substantial part of literature relevant for the simulation of volcanic emission plumes. There have been many studies, which highlight important aspects discussed in this manuscript, which are not even cited (e.g. Brodovsky et al. 2021 or Quaglia et al. 2023).

We thank the reviewer for the comment. We included and discussed more model literature (including the recommended ones) in the introduction of the revised version, and we agree with the reviewer that this literature is crucial to discuss:

"Volcanic SO_2 emission databases are the basis for the correct implementation of volcanic eruptions in atmospheric models. Timmreck et al. (2018) recommend four different emission inventories. Neely III and Schmidt (2016) and Mills et al. (2016) provide an inventory for tropospheric and stratospheric volcanoes, covering daily emissions and providing plume top and minimal height. Carn et al. (2017) provide a very detailed list of tropospheric and stratospheric volcanoes, however not distinguishing between the tropospheric and stratospheric part of the plume. Brühl et al. (2018) developed a volcanic SO_2 emission database from 1990 to 2021 (updated by Schallock et al., 2023), focusing only on the stratospheric part of the plume, also including smaller eruptions. Finally, the emission database from Diehl et al. (2012) only covers volcanoes up to 2010.

The treatment of SO_2 from volcanic eruptions based on the available emission inventories in global atmospheric models varies widely (e. g. Quaglia et al., 2023; Timmreck et al., 2018). Quaglia et al. (2023) performed a model intercomparison study focusing on the Pinatubo eruption, finding that inter- and intra-model differences in the response of SO_2 and sulfur aerosol on the Pinatubo eruption are large for a range of sensitivity experiments. The differences were mostly attributed to differing stratospheric transport, emission databases, aerosol microphysics and stratospheric chemistry. Vattioni et al. (2024) investigated the impact of microphysical settings within the SOCOL-AERv2 aerosol chemistry-climate model and found that the microphysical timestep, as well as the order of the microphysical processes can lead to vastly differing results in stratospheric aerosol burden.

Brodowsky et al. (2021) studied small- and medium-sized volcanic eruptions from 2008 to 2012, investigating the sensitivity of the resulting aerosol burdens to the different emission databases, internal model variability, dynamic nudging and the vertical resolution. The largest uncertainties resulted from the emission databases and their application. Most volcanic model studies inject SO_2 in columns at the geographical location of the volcano (e. g. Brodowsky et al., 2021; Mills et al., 2016; Quaglia et al., 2023; Schmidt et al., 2018), however the vertical extent

of the column is mostly unknown and depends on a number of assumptions. For instance, Brodowsky et al. (2021) emitted the recommended SO_2 burdens for the emission inventories of Brühl et al., 2018; Carn et al., 2017; Diehl et al., 2012 from the prescribed plume top height down a third of the distance to the Earth's surface. However, this approach highly depends on the emission inventory derivation, i. e. if only the stratospheric plume is considered for the derived SO_2 burden (as in the database from Brühl et al. (2018)) or the tropospheric part of the plume as well (as in the databases from Carn et al. (2017) and Diehl et al. (2012)). "

Usually, the 1991 Mt. Pinatubo eruption is used as a basis for model validation (e.g., Quaglia et al. 2023, which even includes EMAC simulations). Why did you not perform a simulation of Mt. Pinatubo to validate your model? Probably because this has been done extensively already, but it would allow comparing also with other models.

We chose the Nabro volcano, as there are extensive SO_2 satellite observations available on this eruption, including 3-D information from MIPAS. This is not the case for the Pinatubo eruption, and there are large uncertainties in observations. We added an explanation on why we chose Nabro: "As it was observed by a number of satellite instruments, it offers a perfect case study to investigate the spatio-temporal evolution of volcanic SO_2 in the stratosphere."

We also want to clarify, that this study is not a specific study on the impacts of volcanic eruptions. We want to present the new submodel, evaluate it, and provide a default setup to achieve realistic SO_2 emission plumes for the last three decades. We do not concentrate on case studies and their atmospheric consequences, such as the Pinatubo eruption; instead we include a large number of eruption plumes reaching the stratosphere. The goal is to provide a standard setup that includes reliable SO_2 emissions also for studies not focusing on volcanic eruptions, especially in the stratosphere region. The extreme event of the Pinatubo did not appear to be a suitable candidate for this evaluation.

The introduction could focus more on topics relevant to the rest of the manuscript. E.g. why do you explain ozone chemistry and ash in such detail if it is not relevant to the manuscript? The manuscript provides a very broad discussion about the general impact of volcanic eruptions on climate and atmospheric chemistry, but it does not discuss the problems and current limitations when modelling volcanic eruptions (e.g., spatial & temporal resolution, model agreement/disagreement of past model studies among each other and with observations, sectional versus modal microphysics modules, previous studies which addressed the vertical/horizontal distribution of volcanic emission plumes). Since you submit to "Geoscientific model development" I think a more technical motivation/introduction would be appropriate. I think the introduction should clearly motivate the research so that it is clear to a reader why the submodule was developed and why there is need for the research provided. In my point of view this is not the case.

We thank the reviewer for this suggestion. Indeed, we talked too detailed about ash and ozone chemistry. We shortened this considerably. However, we wanted to include the discussion on the importance of ash, as EVER was explicitly designed, such that volcanic ash can be simulated.

We chose to focus more on the general volcanic impacts to motivate why this submodel is needed within MESSy, and not necessarily why it is needed in addition to other available

volcanic emission routines in various models. However, we agree with the reviewer, that a more technical focus would improve the introduction. Thus, we discussed limitations in modelling volcanic eruptions more detailed in the introduction of the revised version. We did motivate the research in the original version by showing, why it is crucial to have this submodel in MESSy, but we now made it clearer, why we did not merely take an existing routine from other models. The technical extension to the introduction can be found as a reply to the comment on the additional literature above.

Your conclusions just summarize what you have done in your work, including a brief recapitulation of some qualitative results with a brief outlook for further potential applications of the sub-model in the end. However, I think the conclusion section should summarize the key findings of the paper (also some quantitative statements, not only qualitative) and put them into a broader context. The conclusions should demonstrate the importance of the paper and convey the larger implications of the paper to the research field. Additionally, I also find it important to address the limitations of the research/model provided (or more specifically: of the presented sub-module) and highlight aspects identified in your work which require further research and development. I don't think these points are provided in the current version of the conclusion section of the manuscript.

Yes, we mostly summarized our work, but in addition, we also summarized our key findings in the submitted manuscript, e. g. "Horizontal position and emission timing were found to have a minor impact on the long-term SO_2 burden in the stratosphere. Nevertheless, these parameters play a crucial role in detailed process studies during the initial weeks after an eruption. Overall, we conclude that simulations of volcanic eruptions can be effectively performed with the help of 3D- and column emissions. The optimal approach depends on the specific use case, with column emissions excelling in the short-term, and similar performance in the long-term" and " [...] , and we advocate its inclusion in simulations using the MESSy framework focusing on the upper troposphere and stratosphere. For very strong eruptions, it may be beneficial to distribute the emissions over multiple horizontal grid boxes and an extended time period, or adjust the vertical plume width, if discrepancies with observations occur".

We agree, that we can put them into some broader context with previous findings and point to some limitations and shortcomings, and did so in detail in the revised Discussion section. Additionally, we outlined further necessary research and development in the revised version. However, a technical paper describing a new submodel, evaluating it, and developing a default setup does not necessarily need "larger implications of the paper to the research field" but rather vital information for the people using the submodel. Larger scientific implications to the research field can result from future studies using the submodel. However, as we indeed produce findings, that are important for the research field, we discuss them in more detail in the conclusion and abstract in the revised version (see also abstract below).

Similarly, the abstract just summarizes what you did in this study, but you do not provide any results or conclusions or broader implications to the research fields in the abstract, which I think is an essential part of the abstract.

As outlined above, this is a technical paper describing a new submodel, evaluating it, and developing a default setup. The conclusion basically is, that it was developed, it works and it can reproduce observations. We added additional findings and implications (see comment

above) in the revised abstract:

" This work documents the operation of a new submodel for tracer emissions from Explosive Volcanic ERuptions (EVER v1.1), developed within the Modular Earth Submodel System (MESSy, version 2.55.1). EVER calculates additional tendencies of gaseous and aerosol tracers based on emission source parameters, aligned to specific sequences of volcanic eruptions or other atmospheric emission sources, allowing for the employment of various vertical emission profiles. We show that volcanic SO_2 plumes can be reasonably reproduced through EVER emissions in numerical simulations with the ECHAM/MESSy Atmospheric Chemistry Model (EMAC), using satellite observations of SO_2 columns and mixing ratios following the explosive eruption of the Nabro volcano (Eritrea) in 2011 and a degassing event of the Kilauea volcano (2018) in Kilauea. Previous volcanic studies showed large variability in stratospheric SO_2 burdens depending on the chosen volcanic emission databases and parameters. Sensitivity studies on SO_2 emissions from the Nabro volcano explore perturbations of the emission source parameters, revealing that emission altitude and the emitted mass above the tropopause are most important for the mid- to long-term evolution of stratospheric SO_2 plumes, while the correct timing and geographical location of the stratospheric entrance is crucial for the short-term plume evolution. We integrate information from a volcanic SO_2 emission inventory, additional satellite observations, and our findings from the sensitivity studies to establish a historical standard setup for volcanic eruptions impacting stratospheric SO_2 from 1990 to 2023, successfully evaluated with satellite observations of stratospheric SO_2 burden and aerosol optical properties. We advocate for this to be a standardized setup in all simulations within the MESSy framework concentrating on the upper troposphere and stratosphere in this period. Further potential applications of EVER involve studies on volcanic ash, wildfires, solar radiation modification, and atmospheric transport processes. "

An important aspect which is missing in the analysis and discussion of your results is the role of chemical loss of SO_2. In addition to dilution and transport, a (maybe even more) important aspect which determines the SO_2 lifetime is chemical SO_2 loss (i.e., mainly oxidation with OH and O_3). How is this represented in MECCA, and how does this influence your modelled results. You also write that for the Kilauea study, you applied "simplified chemistry". What does this mean? Since chemical loss is very important for the SO_2 lifetime you need to discuss how this affects the results in your study. How is the reaction with O_3 and OH represented in MECCA? Were the O_3 and OH fields also nudged to observations like written in the model description (probably not)? How does this influence the modelled results?

We agree with the reviewer that the chemical loss of SO_2 is important, and actually is the reason for sulfur aerosols to form. In the supplement, we provide the full chemical mechanism for both simulations, the stratospheric one and the "simplified chemistry" for the Kilauea simulation. We have also shortly added these information in the revised manuscript.

For the stratospheric setup:

"We employed the Mainz Isoprene Mechanism MIM1; Jöckel et al., 2006; Pöschl et al., 2000 within MECCA (see supplement for detailed mechanism), considering oxidation of SO_2 with OH to form SO_3, subsequently reacting with H_2O to form H_2SO_4."

For the Kilauea setup:

"Oxidation of SO_2 to H_2SO_4 is directly realized via reaction with OH, without producing any intermediates. We do not consider DMS and OCS here (see supplement), leading to a potential underestimation of background maritime SO_2 concentrations."

While we treat SO_2 oxidation with OH in both setups as described in the supplement, we do not consider its oxidation with O_3. This reaction only takes place in the aqueous phase and is not important for the stratosphere (e. g. Kremser et al. (2016)). O_3 and OH are part of the chemical mechanism (see supplement) and are not nudged towards observations. The described nudging affects only the model dynamics using meteorological reanalysis data, which is more clearly described in the revised version (see also specific comment on nudging below).

Another aspect which is not discussed is the importance of aerosol microphysics. You show an in-depth analysis of the SO2 plume evolution and then in chapter 4 you suddenly come up with AOD analysis. However, there is an important step missing in between: Aerosol Microphysics. You write that you are using GMXE as an aerosol microphysics module and represent the aerosols using a modal approach. However, you do not provide any information about resulting aerosol burden, aerosol size distribution ect. You only show AOD and extinction (without indicating the wavelengths under consideration). Thus, the whole aerosol microphysics (which is a very important aspects when simulating volcanic emission plumes) is treated as a black box in this manuscript. In my point of view, it is crucial to also show resulting aerosol burden and compare them with observations as well as which other models (e.g. see Brodowsky et al. 2021).

We agree with the reviewer, that aerosol microphysics is a very important aspect for the modelling of volcanic aerosol, and we provide more detail in the revised version (e. g. the microphysical timestep of 8 minutes). Thank you for pointing this out. However, we want to make clear again, that this manuscript is not aiming to evaluate the aerosol microphysics or study the sensitivity of the results to the aerosol microphysics. Instead, we focused on the initial SO_2 plume, and only used the comparison to OSIRIS for an additional validation, and to compare our results to Schallock et al. (2023).

We thank the reviewer for pointing to Brodowsky et al. (2021), and we apologize for not including it in the initial submission. However, the only decadal satellite observations of (sulfur) aerosol burden available we could find is the MIPAS dataset from Günther et al. (2018), that was also used by Brodowsky et al. (2021). We considered using this dataset. However, a number of assumptions have to be made for a model-observation comparison, including the density of the H_2SO_4-H_2O system, which fraction of the sulfur aerosol is liquid or solid (not directly distinguishable in our model), and how much of the aerosol water to include. The uncertainties in these assumptions lead to large variations in our model. Brodowsky et al. (2021) simplified this by normalizing the comparison to background levels, or adding a baseline correction to MIPAS, not showing the absolute differences. However, we believe that this would not add much benefit to our work, as the results would look very good, but only due to the normalization. However, if it is necessary, we could add this comparison to the supplement.

Moreover, in our opinion our study has a different focus, as we want to evaluate the EVER submodel. Thus, we look into the uncertainties in the SO_2 emission parameters, while not considering uncertainties in aerosol microphysics in great detail. Our manuscript is a "Development and technical paper" in Geoscientific Model Development, while the Brodowsky et al. (2021)

paper is a scientific study in JGR Atmopheres focusing on the evolution of sulfate aerosol, and thus focuses more on the production of H_2SO_4 and the resulting aerosol burden.

We apologize for not mentioning the wavelengths under consideration, and we have added this information in the revised version. The extinction and AOD are considered at a wavelength of 750 nm. Additionally, we will compare our results to previous model results. We decided to focus on comparison to observations of SO_2, as this is mostly influenced by the EVER submodel, and we wanted to show that we can reproduce observed SO_2 mixing ratios. As pointed out, the resulting aerosol burden depends on a lot of different factors, the evaluation of which goes beyond the scope of this manuscript. We clarify in the revised version, that the comparison of simulated AOD to observations is only a validation of the changes compared to the previous study by Schallock et al. (2023), and not a separate study on aerosol formation.

Another aspect which comes too short in the discussion section is the influence of the spatial and temporal resolution of different processes. It is well known that these aspects are very important for realistic representation of volcanic plumes. While for the troposphere the horizontal resolution is more important in the stratosphere the vertical resolution is more important. For aerosol microphysics the microphysical timestep should be set small enough to realistically simulate nucleation and condensation. You could mention these aspects in the discussion of your results.

Yes, the spatial resolution is very important for the modelling of volcanic eruptions. For that reason, we chose the maximum available horizontal resolution for the tropospheric study of Kilauea (T255), and the maximum available vertical resolution (90 levels) for the stratospheric study, similarly to the work of Brodowsky et al. (2021). The evaluation of different resolutions again goes beyond the scope of the paper.

We do not distinguish the chemical and microphysical timestep in our simulation, as the general model timestep length is only 8 minutes (we have added this information to the revised manuscript), and chemistry and microphysics are calculated in each timestep. This is different to the recent publication by the reviewer (Vattioni et al., 2024) where chemical and microphysical timestep differ. Vattioni et al. (2024) showed that in volcanic simulations, first considering condensation and nucleation results in the smallest numerical error. This is the default in GMXe, and thus we did not discuss it. Further reducing the timestep will of course improve the results, but is not computationally feasible. We further argue, that the chemical timestep length of 2 hours in the work from Vattioni et al. (2024) will lead to different errors in microphysics compared to the 8 minute chemical timestep length in our EMAC simulations. Moreover, the main focus of the manuscript is on the initial SO_2 plume, that is not influenced by the microphysical settings. For that reason, we do not want to engage in this discussion.

I also find the manuscript too long. Many aspects which are discussed in the introduction and submodel description are not relevant to the storyline or are not picked up again in the discussion. I suggest shortening substantially and putting part of the text (e.g. description of code and namelists) into the supplement. Also, the structure could be improved. For example, the model description and setup and description of the observations (sect. 3.1 and sect. 3.2) could be a chapter for its own or part of chapter 2. Because the setup described there is also used in chapter 4. The different events simulated here (i.e., Nabro, Kilauea and the

2007-2011 period) seem a little disconnected to each other. Why don't you show a full analysis of only one event (e.g. Nabro), but in more detail including sulfuric acid aerosol burden etc.

We thank the reviewer for the very helpful suggestions, that we applied to our manuscript. First, we considerably shortened the manuscript. The main affected Sections are described below:

- The description of general impacts of volcanic eruptions were shortened in the introduction.

- We removed the namelist example for the new submodel (Fig. 1) and some of the detailed explanations, and moved it to the supplement.

- We shortened the introduction of the Nabro and Kilauea volcano.

- We restructured the methods, and thereby avoided repetitions.

In addition, we generally restructured the methods and results section. Sect. 2, including the presentation of the EVER submodel mostly remained in the same structure, only adding the description of the default namelist setup as an additional subsection. Sect. 3 now only focuses on the methods and the experimental setup, while Sect. 4 only shows the results of the three experiments. This results in the following new structure for Sect. 2-4:

2. New MESSy submodel for Explosive Volcanic ERuptions (EVER)

   2.1. Submodel description

      2.1.1. Primary aerosol emissions

      2.1.2. Vertical distributions

   2.2. Historic default namelist setup

3. Methods

   3.1. Observations

   3.2. Model setup and numerical experiments

      3.2.1. SO_2 in explosive volcanic eruptions — Nabro (2011)

      3.2.2. Multi-year simulation with the historic namelist setup

      3.2.3. SO_2 from degassing volcanoes — Kilauea (2018)

4. Results

   4.1. SO_2 in explosive volcanic eruptions — Nabro (2011)

   4.2. Evaluation of the historic default namelist setup in a multi-year simulation from 2008 to 2011 (stratospheric SO_2 burden and sAOD)

   4.3. SO_2 from degassing volcanoes — Kilauea (2018)

However, we wanted to keep the submodel description in a separate section, as this is a very important part of the manuscript. Moreover, we do not believe that the submodel description

in a technical paper has to necessarily follow the storyline of the evaluation, but should instead focus on completeness.

We understand, that the different events seem to be disconnected. We motivated the methodology in more detail in the experimental setup section in the revised version. We chose the different events for the following different reason:

- Single analysis of the Nabro volcano to evaluate the submodel in the stratosphere, and perform sensitivity studies on the emission parameters for stratospheric volcanoes.

- Kilauea study to show the applicability of the submodel for degassing volcanoes

- 2008-2011 period to evaluate the historic namelist setup developed

For that reason, we want to keep the different evaluations, with an improved structure, in the revised version.

Detailed comments

Line1: What is a methodological study? Isn't every scientific study methodological?

We agree and removed the word "methodological".

Line 14/15: Suggesting to change "solar geoengineering" to "solar radiation modification", a more appropriate term.

We replaced "solar geoengineering" by "solar radiation modification" in the abstract, and all later occurrences.

General Comment on the abstract: The abstract mostly reflects what was done in this study, but there is no mentioning on results and conclusions, which I think is a key component of an abstract. Thus, I suggest adding some quantitative results and conclusions/broader impacts.

We slightly extended the abstract, including the findings of our study that are important for the community. The full abstract of the revised version can be found above.

Line 19/20: "On the on hand... on the other hand" is normally used for opposing arguments. The ones mentioned here are more additive. I suggest reformulating. Whether it is "substantial" or not, depends on the magnitude.

We reformulated that, thank you for the remark: "The resulting additional stratospheric aerosol loading may exert a negative radiative forcing (Schallock et al., 2023; Schmidt et al., 2018), and serves as surfaces for heterogeneous reactions, thus, impacting stratospheric composition in general and ozone in particular (Klobas et al., 2017; Tie and Brasseur, 1995)."

Line 25 and lines 32-34: Do you have references for that?

The source for line 25 was also Textor et al. (2004) as the following sentence. However, we removed the sentence anyways, as it is was not important for the introduction, and we aim to shorten it, following the general comments above.

Line 25/26: Do you have a reference for this?: "The composition of volcanic plumes exhibits considerable variability and depends on the intricate mixture of chemical species in the magma"

The source was also Textor et al. (2004) as the following sentence. However, we removed the sentence anyways, as it is was not important for the introduction, and we aim to shorten it.

Line 27-31: I suggest combining the two sentences (i.e., list the example of Hunga Tonga in the first sentence).

Thank you for the suggestion. We reformulated it accordingly.

Line 32-34: a definition of "long-term" would be good. sulfuric acid aerosols and their precursors are usually removed from the stratosphere within 2 years. I don't think this is long-term in terms of climate. Maybe chlorine species would have a long term effect.

We removed "long-term".

Line 34-36: This is somehow confusing. In the first paragraph you speak of "the most explosive volcanic eruptions" and now you write of emissions per year. Do you still speak of large explosive volcanic eruptions or do these numbers also account for degassing non-volcanic eruptions? I suggest being more precise here to what exactly these numbers refer to.

Thank you for the remark. We removed this part as it was not important for the introduction.

Line 37/38: I suggest changing the term "sulphate" with "sulfuric acid", since technically speaking, a sulphate is a solid (e.g. CaSO4). Also add references to this statement.

We used the term sulfate, as it is common practice in the literature, but we agree with the reviewer, that it is imprecise. However, we prefer to use the term "sulfur aerosol" over "sulfuric acid aerosol", as it also includes (bi-)sulfates, such as ammonium (bi-)sulfate, that can form in the presence of ammonia. We adjusted it everywhere. We also added a reference (Kremser et al., 2016) to the mentioned statement.

Line 36: Maybe add "under volcanic conditions", otherwise other sulfuric acid precursor gases are also important. I know SO_2 is poisonous, but isn't the effect on acid rain mainly a result of uptake of sulfuric acid aerosols (and not primarily SO_2)?

We removed this passage in the process of shortening the introduction. We agree with the reviewer, that acid rain is a result of uptake of sulfur aerosol, and not SO_2. We implied that "SO_2 plays a pivotal role in air pollution and the formation of acid rain" via the formation of sulfur aerosol and subsequent uptake, but did not explicitly write it.

Line 46: "up to" or "over"? But not both.

Removed "up to", and reformulated to "The lifetime of stratospheric aerosols can exceed 2 years when injected in the tropics".

Line 81: What is a "horizontal gid box"? What is the difference between a horizontal or vertical grid box? Do you mean "vertical column of gird boxes"?

Yes, that was the intended meaning. This part was majorly reworked, when giving a more technical introduction, and this formulation does not appear any more in the revised version (see also the replies to the general comments on the more technical introduction).

Line 85: Are you only aiming at providing or are you providing? I suggest reformulating to "We provide . . . ", or reformulate in another way if you don't. Same for the last sentence of the paragraph: "This was achieved through the following three steps of work: ". Don't undersell your research.

Thank you for the suggestion, we majorly reformulated the respective paragraphs including the suggestions.

Line 90: ". . . of vertical emission distributions . . . "

As we specified, that we talk about emissions in the sentence already, we thought it was not necessary here. However, we added "emission" to make it clear.

Line 116: You only use GMXe aerosol microphysics module in this study. Why do you introduce the other two submodules too? This only causes confusion.

The new submodel EVER was developed, such that it can be coupled to the different aerosol submodels in the MESSy environment, and the technical functionality was tested for all three submodels. We wanted to make that clear at that point, as this is an essential part of the submodel. We added an explanation, that primary emissions were not explicitly evaluated, and only GMXe was used in all studies to process the resulting sulfur aerosol. Additionally, we moved this paragraph to the Subsection on "Primary aerosol emissions"

Line 127: What do you mean by "linear columns"?

We reformulated to "... capable of emitting point sources and uniformly distributed columns of trace gases". We replaced all occurrences of "linear columns".

Section 2.1: I think the description of the new submodel is too technical. It is probably not useful for most readers of this study. I suggest making the description of the new sub model more general and provide a more technical description (e.g. how the name list works, what the different name list parameters are) in the supplement.

We thank the reviewer for the suggestion. We moved the namelist to the supplement, however kept the explanation for the different parameters, that need to be provided, as we believe this is essential for a documentation of a new submodel.

Section 2.2.1: The title of this section is "Primary emissions", but the subchapter is specifically on direct aerosol emissions. Maybe specify this in the title. However, why is this subchapter important? In this manuscript, only SO_2 injections are evaluated. Maybe you can skip this subchapter or put it into the supplement.

We changed the title of the section to "Primary aerosol emissions". This work documents the new submodel, which is capable of directly emitting aerosols. We only performed the sensitivity studies and evaluation for SO_2. However, we believe that it is important to keep this paragraph

for completeness, as the capability for primary aerosol emissions is essential for the submodel, and its documentation for the people using it.

Section 2.2.2: This could be picked up later in the paper.

We believe that the reviewer refers to Section 2.1.2, as there is no Section 2.2.2. We pick this up later in the experimental setup (Gaussian distribution) and in the discussion (distribution of emissions on a larger area).

Section 3.1 There is no mentioning of the microphysical, chemical and dynamical time steps applied in the models used in this work. A recent study has highlighted the need for appropriately setting the microphysical time step when simulating volcanic eruptions.

We thank the reviewer for this comment, and provided the model timestep (8 minutes) in the revised version (identical for microphysics, chemistry and dynamics), and also refer to Vattioni et al. (2024) in the introduction.

Section 3.2: Maybe this section can be shortened. Are such detailed descriptions of all the different satellite and technical details such as their resolutions required?

We thank the reviewer for the suggestion. We considerably shortened this paragraph, by providing a table on the used satellites, and refer to the references for more detail. In the new Section on the Experimental Setup (3.2), we give more information on which satellite observations we use for which evaluation. The new table is provided in the reply to Reviewer 2.

Section 3/4: I assume the model description provided in section 3.1 also applies to section 4, right? And some of the described observations (satellites) in 3.2 are only used in section 4, right? To avoid confusion, I suggest separating the model description and observations from chapter 3 and create an own chapter for this. Or maybe something similar, just improve the structure of the paper, it is confusing sometimes.

We agree with the referee, that the chosen structure can be confusing, and are thankful for the suggestion. Indeed, the model description and observations of Sections 3.1 and 3.2 also applied for Section 4. We rearranged the structure, providing a separate Section for the methods, where we introduce the satellites, the model setup and outline the experimental setup. We also made it clear, which model setup and which observations contributed to which experiment. See also the reply to the last general comment for more details on the applied major changes to the manuscript structure.

Line 203-207: Why is this information important at all? You are not looking at aerosols in this chapter, but only at SO_2 plumes. Is SO_2 also treated within GMXe?

We agree, that GMXe is not important for Section 3, as SO_2 is not treated with GMXe. We wanted to give an overview on the general model setup, that was used throughout the manuscript, but understand that the structure was confusing and adjusted it (see previous comment). The microphysics are now only introduced for the multi-year simulation.

Line 200: Nudged to which variables? Wind and temperature?

We reformulated to:

"Temperature, the logarithm of the surface pressure, divergence and vorticity are "nudged" (more details in Jöckel et al., 2006) towards meteorological reanalysis data (ERA5, Hersbach et al., 2020) from the European Centre for Medium-Range weather forecasts (ECMWF). In addition, we employ the QBO submodel to weakly nudge the simulations to QBO zonal wind observations between 10 and 90 hPa (Giorgetta et al., 2002) to avoid a phase drift."

Line 222: Maybe just write: "Namelist setup, chemical mechanism and runscripts can be found in the supplement." However, I think this belongs in the data availability statement not in the main text.

We rephrased it accordingly. However, we want to provide the information here, so the reader can refer to the supplement, in case of any question to the chemical mechanism.

Line 216-219: Why is this information important? In this chapter (chapter 3) you are only focusing on SO_2 plumes, but no aerosol optical effects. I suggest skipping.

We agree with the reviewer. As mentioned above, we rearranged the manuscript, and the aerosol optical properties are now part of the experimental setup section, and only introduced for the outline of the evaluation of the default namelist setup.

Line 284/285: "Especially the second stratospheric plume on June 16 could comprise remnants of the tropospheric plume, that are uplifted" This sentence is confusing to me: What do you mean with tropospheric plume? The volcanic plume or the monsoon?

In the preceding paragraph (Lines 276-278), we discuss the separation of the tropospheric and stratospheric plume: "Concurrently, a wind shear within the AMA around the tropical tropopause induced a separation between tropospheric and stratospheric SO_2, with the stratospheric component further north." So we speculate (or previous studies speculated) if the second stratospheric emission plume actually results from the uplift of the tropospheric plume from June 13. We clarify in the revised manuscript, that we refer to the previous plume.

Line 323: There is no specification of the emission in Mills et al. 2016 so far. This pups up here a little abruptly, since this was not discussed in the introduction or anywhere prior to here.

We agree, and we discuss different emission inventories in the introduction and Mills et al. (2016) in particular in the revised version (see above in the replies to the general comments). In addition, we set the different emission inventories into perspective, and clarify, why we decided for Schallock et al. (2023) in the Discussion section of the revised manuscript. Moreover, we give more information about the Nabro emissions from Mills et al. (2016) in the revised manuscript: "Emissions as described by Mills et al. (2016). The emissions are injected over several days and in columns covering the free troposphere and the UTLS region, ranging from 2.5 to 17 km. The only partly stratospheric emission takes place on June 13, emitting 1500 kt, uniformly distributed over the altitude range from 9.7 to 17 km."

Line 334: "The column amount estimation assumes that all SO_2 of the plume is

centered at the respective altitude depicted in Fig. 4 " To me it is not clear how this explanation should explain the discrepancies between observed and modelled column amounts. Can you explain further?

We agree, that this statement does not sufficiently explain the discrepancies, as it was not formulated precisely. In fact, the retrieval considers that the plume is centered at these altitudes above 14 km. However, it takes into account the whole column for this retrieval, also including the tropospheric part of the plume, thereby not only representing the stratospheric part. We added the following sentence after the cited sentence: "However, it takes into account the complete column, also including the tropospheric part of the plume."

Line 338: "From Fig. 4, it seems that the simulated columns slightly broaden over time compared to the observations, with the plume appearing to sink." I am confused here. Figure 4 shows altitudes not column SO_2. It also seems like the simulated plume (reference) is higher up compared to the observed plume. Do you mean the observed plume appears to sink? It is hard to see any broadening of the plume in Figure 4.

We want to shortly clarify the two statements in the following:

- "... it seems that the simulated columns slightly broaden over time compared to the observations ...": We wanted to explain here, why in the bottom row of Fig. 4 (June 17) the simulated altitude distribution is broader than in the observations. The reason is that we see maximum SO_2 mixing ratios in the model at these altitudes, however at the edges of the plume these values are very small, below the detection limit of IASI, which can be seen from Fig. 5 bottom row, that shows a much narrower plume.

- "... with the plume appearing to sink.": This refers again to the bottom column of Fig. 4 (June 17). At longitudes higher than 90° E the reference simulation shows lower altitudes compared to the observations. This is, what we wanted to comment on.

To make the first statement clearer, we reformulated as follows: "While the simulated altitude distribution broadens over time (Fig. 4, June 17 **reference**), the total column analysis (Fig. 5, June 17 **reference**) shows that the columns at the edges of the plume are very low, falling below the detection limit of IASI." We omitted the comment on the sinking to avoid confusion.

Line 339/340: The simulated (reference) SO_2 column distribution in Figure 5 is broader compared to the observation. . . not narrower. This is confusion.

See reply above.

Line 361/362: "However, IASI faces limitations in capturing the long-term evolution of volcanic plumes due to the dilution of the emitted SO_2, leading to column amounts that fall below the instrument's detection limit" Do you really know dilution is the main process that SO_2 concentrations fall below the detection limit? If yes, do you have references for this? Isn't chemical loss (SO_2 oxidation via OH and O3) equally or even more important on longer time scales? How is this represented in the model and how does this affect the long-term evolution of the SO_2 plume?

We moved this statement to the Methods section, and reformulated to. "This can likely be

attributed to the dilution and chemical removal of SO_2 in the plume." We talk about the removal in the Experimental setup in more detail (see general comments).

Lines 416-422: "The overall slightly faster decline observed in the simulation compared to the observations may be a consequence of the absence of primary particles, such as volcanic ash, in the simulations, resulting in a discrepancy between simulated and observed particle size distributions."
What processes should be the reasons for that? I guess you mean that ash could result in self-lofting of airmasses due to absorption of radiation and thus local heating? Or what other processes do you have in mind? It is important to name them since this is not clear from how it is written now.

"Alternatively, the simulated particle sizes may grow excessively large too quickly, leading to an overestimation of sedimentation efficiency"
You do not show any simulated particle sizes. You show and write about SO_2 plumes. SO_2 is a gas. Gases are mainly subject to diffusion & transport and in the case of SO_2 more importantly: chemical loss, ... but definitely not sedimentation. What about chemical loss? How does this affect the dissipation of the plume compared with observations?

"Whether this discrepancy arises from nucleation rates versus condensation efficiency, the overall representation of the size distribution with only four modes, or the limitation to one horizontal grid box will be the topic of upcoming studies."
SO_2 concentrations are definitely not affected by nucleation and condensation rates. Chemical loss of SO_2 is dominated by reaction with OH and O3. These reaction result in formation of SO3, which then together with H2O forms H2OS4 gas. H2SO4 gas has a very low vapor pressure and immediately forms sulfuric acid aerosols via condensation or nucleation. Have a look at Feinberg et al. 2019 and the stratospheric sulfur cycle presented in there. It should get obvious that nucleation and condensation rates as well as aerosol size distributions do not affect the chemical SO_2 lifetime/burden.

Feinberg, A., Sukhodolov, T., Luo, B.-P., Rozanov, E., Winkel, L. H. E., Peter, T., and Stenke, A.: Improved tropospheric and stratospheric sulfur cycle in the aerosol– chemistry–climate model SOCOL-AERv2, Geosci. Model Dev., 12, 3863–3887, https://doi.org/10.5194/gmd-12-3863-2019, 2019.

We agree, that the discussion in this paragraph was not correct, as we attributed the decline of the stratospheric SO_2 burden to aerosol effects, while they are actually not affected by these. We replaced the paragraph by the following in the revised version:

" The overall slightly faster decline of the stratospheric SO_2 burden in the simulation compared to the observations appears consistent across all simulations. It can either be attributed to an overestimation of the chemical removal, i. e. the oxidation with OH, or too efficient transport from the stratosphere to the troposphere, that also depends on the injection altitude. "

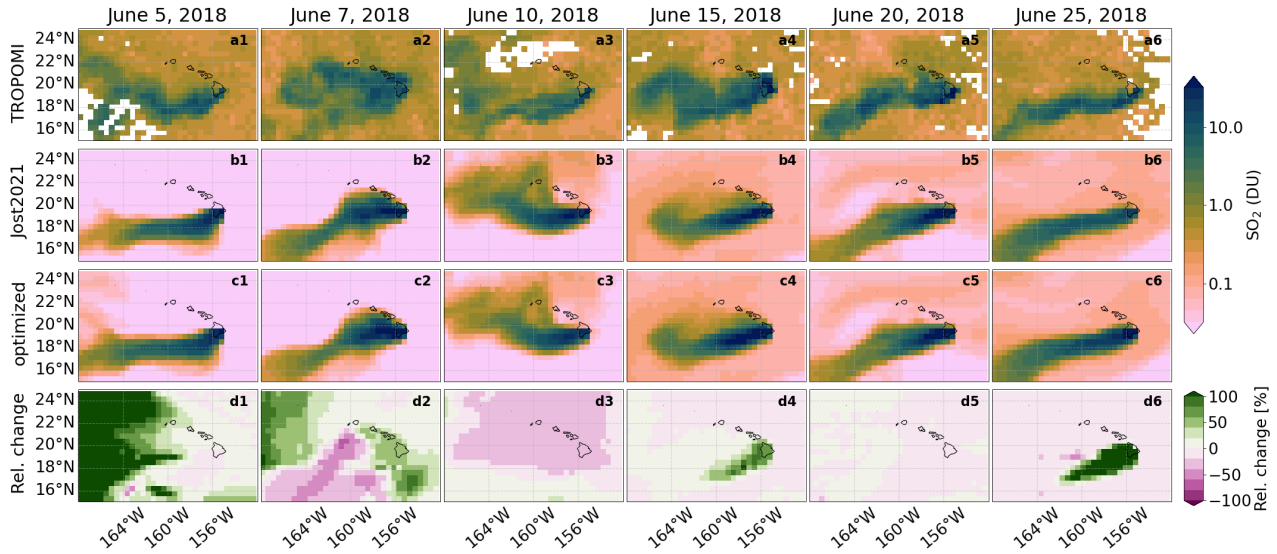Line 441: What "data"? Simulated or observed?

Figure 1: Observed (top) and simulated SO$_2$ columns resulting from the degassing of the Kilauea volcano at selected days in June 2018 at a model resolution of T255. **a*,** Observations from TROPOMI, regridded to T255 resolution. **b*,** Simulation with emission rates derived by Jost (2021, scaled by a factor of 4.3 – see text for details). **c*,** Simulation with optimized emission rates, based on a comparison between model and observations (refer to the text for more details). **d*,** Relative change from the simulation with derived emission rates by Jost (2021) to the optimized simulation.

We believe that the reviewer refers to line 341. We refer to the simulation here, and state this in the revised version. We apologize for the incorrect formulation

Line 341/342: Ahaa... it only becomes clear that you were talking about the initial plume on June 14 until now. The statements you make in this paragraph are only true for the initial plume on June 14. You really need to be more precise here... The statements of this paragraph are not valid for June 17.

We actually refer to June 17, where the altitude distribution is broader in the simulation compared to observations. We reformulated as mentioned in the reply to comments above.

Figures: I suggest assigning letters a, b, c, d ... to the subpanels of figures to enable better referencing.

We thank the reviewer for the suggestions, and applied letters to all manuscript figures with more than one panel (see also updated figures in this reply).

Figure 444: It is hard to see any difference in agreement/disagreement with observation in Figure 8 of the June 7 and 10 data compared to for example June 5 and June 15. I suggest plotting the differences compared to the observations in the middle and lower panel. This would highlight the differences.

We believe that the reviewer referred to Figure 8. We adjusted the Figure and also added the relative change between the simulations (see Fig. 1d here).

19

Line 448-451: Most of this can go into the figure caption.

We thank the reviewer for the suggestion. We shortened it here, as the information was already present in the caption. The shortened paragraph is (note that the figure numbers changed as the manuscript was restructured):

" Figure 10 provides a quantitative assessment of the horizontal extent depicted in Figure 9 (spanning from $-168°$ to $-152°$ W, $15°$ to $25°$ N), showing spatially averaged SO_2 column amount within this horizontal window, $\overline{SO_{2(col,d)}}$ as observed and simulated at each day in June 2018 d (bottom), and spatial correlation between simulated and observed logarithmic SO_2 column amounts within this window (top). "

Line 556: I disagree with that. The observations and the model does not "exhibit similar patterns" between 0 and 25N. The QBO signal is much more pronounced in the model compared to the observations. And the observed extinction is very different compared to the modelled ones in absolute numbers.

We thank the reviewer for this comment. Indeed, the QBO signal is more enhanced above 20 km, however we mostly focused on the lower stratosphere. Also, the absolute mixing ratios differ, but this is mostly in line with the finding from the sAOD study, or due to the larger vertical extent of the observed volcanic plumes, that are typically represented in AKMs, that were not available for the OSIRIS observations. We reformulated the paragraph to be more precise:

" Figure 8 displays the aerosol extinction at 750 nm, as observed from the OSIRIS instrument and simulated. The effect of the three major volcanic eruptions is mostly evident between 16 and 20 km in the tropics, and between 12 and 18 km at higher latitudes. Analogously to Fig. 7, the magnitude of extinction is differing between the observations and the two simulations. In addition, discrepancies are noticeable in the maximum altitude of the plume and below the tropopause. While the 3D simulation largely reproduces the observed maximum altitude, using satellite observations with, in case of MIPAS, the incorporated AKM as input, the EVER simulation may present more realistic maximum altitudes. The differences between the observations and the simulation below the tropopause are strongly driven by the coincidence with clouds which hinder the retrieval of aerosol extinction. The slight differences between the two simulations from 10 to 15 km in the tropics can most likely be attributed to the differences between the simulation setups. In addition, the QBO signal with differing aerosol concentrations above 20 km, depending on the QBO phase (e. g. Hommel et al., 2015), is more pronounced in the simulation, potentially subject to future studies. "

Line 459: Why do you think the observations are wrong? It could well be your model which is wrong. Why don't you optimize your model to improve agreement with observations?

We do not state, that the observations are wrong. We optimize the emission rates, such that the model can reproduce the observations. So we do not "optimize" the observations, but the input to our model.

Line 460: With "implemented emission rates" you mean the observations, right?

We mean the emission rates, that were derived by Jost (2021) based on observed SO_2 columns.

We do not argue, that these emission rates are wrong, but that they produce wrong SO_2 columns in the model. Thus, we want to optimize them, such that the model produces observed SO_2 columns.

We state in lines 433ff how the emission rates were calculated based on observations. These emission rates were used here.

No, the fit is done as described in the text. $\overline{SO_{2(\texttt{col},d)}}$ is the simulated spatially averaged SO_2 column amount within the depicted horizontal window. Thus, we get a relationship between emission rates of the preceding three days and the resulting simulated average column. Using this relationship, we then use an SGD algorithm to optimize the emission rates, so that we match the observed spatially averaged SO_2 column amount. So yes, it is more or less obvious, that we match the observations in the end. However, that was the goal to show a method, how observations can be reproduced with the model.

We explain it in more detail and give a reference in the Methods section:

" In addition, we perform a simulation with tuned emission rates. For that purpose, we first fit the emission rates from Jost (2021) to the resulting simulated SO_2 columns. Subsequently, we optimize the emission rates gradually by using a stochastic gradient descent algorithm (e. g. Ruder, 2016) to match the observed columns. More details on the optimization can be found in the supplement. "

The algorithm was moved to the supplement.

We wanted to test, if we see improvements in the spatial correlation, as the "correct" representation of the temporal emission over several days may result in the right spatial pattern. However, this was not the case, as the observed SO_2 columns were mostly driven by the emissions of the previous 24 hours.

We only stated, that the model can "effectively" simulate this. We remove the word "effectively". However, we do not claim to do it better than other (regional) models (with higher resolution).

Line 474-478: You did not investigate any sensitivity to spatial resolution. Thus, you cannot make this conclusions. Delete this part, or show evidence for this conclusions.

We did this sensitivity study by performing the simulation in different horizontal resolutions, but did not show it here to avoid extending the manuscript even further. We slightly reformulated the sentence to make it clearer, that we performed these studies:

" Simulations with equal emissions performed at more standard horizontal resolutions, such as T63 and T106, failed to reproduce the observations adequately (not shown), whereas these resolutions are mostly sufficient for stratospheric simulations. "

Line 477/478: This is the most critical result which I think you must discuss more. You only get good agreement with total column SO_2 observations if you tune the emission rates according to your simulation results. I know that this is a common problem for models simulating volcanic eruptions (e.g. Mt. Pinatubo), but you should highlight this. critically discuss it and derive the right conclusions. I also think "analysis" is not the right word here. More precise would be "tuning".

We moved this part to the discussion, used the word "tuning", and clarified that we explicitly use the observations for that. However, we do not see a problem with this approach. We even make clear in Line 636-639, that this could be a process how to incorporate SO_2 from degassing volcanoes into the model, if observed values should be reproduced. We do not state, that the optimized emission rates are determined independent of observations or modelling.

Line 520/521: The magnitude of the signal would not change if the satellite signals were only delayed compared to observations. Isn't it mainly the sensitivity of the measured satellite signal? Please be more precise here.

We wanted to state here, that the overestimation of the model can be attributed to the limited reliability of MIPAS observations directly after the eruption. We agree, that the reason for that is the sensitivity of the measured satellite signal. We slightly reformulated this to: "This discrepancy can be attributed to the limited reliability of MIPAS observations shortly after strong eruptions (see Sect. ..."

Line 522-529: Here again: What is the impact of chemical loss of SO_2? Also did you compare your model to background sulfur cycle (see Brodovsky et al. 2024)? It might make sense to fist perform same simulation as in this study to compare to other models and observations.

Brodowsky, C. V., Sukhodolov, T., Chiodo, G., Aquila, V., Bekki, S., Dhomse, S. S., Höpfner, M., Laakso, A., Mann, G. W., Niemeier, U., Pitari, G., Quaglia, I., Rozanov, E., Schmidt, A., Sekiya, T., Tilmes, S., Timmreck, C., Vattioni, S., Visioni, D., Yu, P., Zhu, Y., and Peter, T.: Analysis of the global atmospheric background sulfur budget in a multi- model framework, Atmos. Chem. Phys., 24, 5513–5548, https://doi.org/10.5194/acp-24- 5513-2024, 2024.

The model was evaluated for background sulfur before in Brühl et al. (2012). The suggested

comparison would be a new study, that could be performed in the future. We additionally now show, that this underestimation could additionally depend on the lower integration limit for the derivation of the stratospheric SO_2 burden (see Fig. 2), and is mostly related to smaller volcanic eruptions, where we did not optimize the stratospheric entry point from IASI, as observations were not available. We adjusted this paragraph in the revised manuscript.

Line 530: Here you suddenly start talking and comparing AOD resulting from these volcanic eruptions. So far you talked and compared SO_2 plumes. It would be great to first see some sulfuric acid aerosol size distribution or how the sulfuric acid aerosol plume/burden evolves in the aftermath of these volcanic eruptions (see Brodowsky 2021). This is what defines the AOD downstream not the SO_2 plume. There is an important part missing here when going from SO_2 plumes to AOD. Without this intermediate step it is hard to say where the discrepancies between model and observations are coming from. It is just guessing since aerosol formation and distribution in the model appear like a black box to the reader... Also, crucial information is missing about the wavelengths to which the AOD and extinctions shown in Figure 11 and 12 are referring to. Why did you only look at 0° to 25°N and 45°-80°N? and not other regions?

We agree that it would be interesting to study the size distributions or the aerosol plume evolution. However, this was not the focus of this manuscript. We still mostly focus on SO_2 plume evolution. We just additionally added the aerosol effects for a short validation and comparison to Schallock et al. (2023). We agree that it is hard to say where the discrepancies are coming from. However, we believe that a more detailed study on this would be outside of the scope of the manuscript. We clarified in the manuscript, why we only show AOD and not all aerosol properties, and provide reasons for not showing comparisons to observations of aerosol burdens in the replies to the general comments. Additionally, we clarify our approach in the experimental setup section revised manuscript.

We apologize for not providing the wavelengths for extinction and AOD, but will add it in the revised version (750 nm). We only looked into the mentioned regions, as (1) all major volcanic eruptions happened in the Northern Hemisphere, and (2) the tropopause has a steep altitude gradient in the region between 25° and 45° making a comparison of sAOD to observations more challenging and less reliable, as it is difficult to diagnose the troposphere in the observations.

Line 537-544: This paragraph needs references and is somehow handwaving.

See comments below.

Line 539/541: I think this reads a little hand wavy here. Please be more specific. A paper which addresses some potential effects is Vattioni et al. 2024. Vattioni, S., Stenke, A., Luo, B., Chiodo, G., Sukhodolov, T., Wunderlin, E., and Peter, T.: Importance of microphysical settings for climate forcing by stratospheric SO_2 injections as modeled by SOCOL-AERv2, Geosci. Model Dev., 17, 4181–4197, https://doi.org/10.5194/gmd-17-4181-2024, 2024.

We agree, that this speculation is not well justified. As the reviewer correctly pointed out, that we cannot really make assumptions about the microphysics, because we do not explicitly

investigate the microphysics, we removed this part, and replaced it (see above).

Line 542-544: "This phenomenon could potentially be addressed by distributing emissions across multiple horizontal grid boxes and releasing the SO_2 over an extended time period." Why would you do this? In section 3 you showed good spatial agreement with observations, so why change the spatial distribution? What I think could help might be changing the horizontal resolution. It seems you again are looking for the error in the emission scheme/observations instead of in within the model. Your suggestion would reduce the SO_2 concentrations and thus the H2SO4 concentrations downstream. This reduces condensation and especially aerosol nucleation rates. But why should this be justified?

See above.

Line 546/547: "This anomaly could be attributed to an overestimation of transport from higher latitudes to the tropical stratosphere, or a general overestimation of the emissions." Why should this be the case? Isn't the transport in this region going exactly into the other direction (from the tropics to higher latitudes)? And there is also a tropical "transport barrier".

Indeed, the transport in the stratosphere follows the Brewer-Dobson circulation and transports air from lower to higher latitudes. However, in the two cases of Sarychev and Kasatochi, the injections happened at higher latitudes (47°N and 50°N, respectively), yet they had an effect on the tropical stratosphere (also in the observations). Thus, there has to be some transport from the higher to the lower latitudes. As both volcanoes erupted in the Asian monsoon period, a possible transport pathway would be the monsoon anticyclone.

Line 549: What differences are you talking about? Difference compared to what?

We mean the discrepancies between the two simulations, and clarified it in the revised version: "The discrepancies between the two simulations in the mid-term transport after ..."

Line 552/553: "The interaction with the South Asian monsoon anticyclone potentially causes differing transport to lower or higher latitudes, respectively." Weren't the simulations nudged towards observed wind fields?

The simulations were nudged to reanalysis data (see above). However, the two simulations differ in altitude, geographical location and timing of the emission. Thus, they are differently affected by the monsoon anticyclone. We extended the statement as follows: "The interaction with the South Asian monsoon anticyclone potentially causes differing transport to lower or higher latitudes, respectively, if altitude, timing and geographical location of the emission do not coincide."

Line 554/555: This sensitivity needs to be addressed by showing some plots in the supplement with different cutoff altitudes, since this defines whether the model agrees with observations or not...

We agree that it is important to show the sensitivity to the cutoff altitude. As the term "stratospheric cutoff altitude" might be misleading, we generally replaced it with "lower integration limit" and made it clear everywhere. Note that this "lower integration limit" refers to the derivation of the stratospheric SO_2 burden and the sAOD, and not to the injection range of
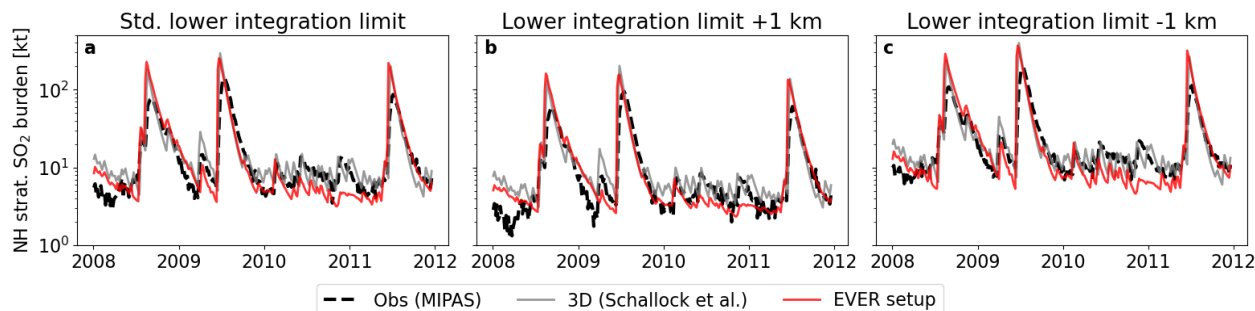
Figure 2: Sensitivity of the evaluation results to the lower integration limit (applied to observations and simulations) for the derivation of the stratospheric SO_2 burden. **a,** Lower integration limit as applied in the manuscript (16 km from 0-30°N, 14 km from 30-60°N and 12 km from 60-90°N. Variations of plus 1 km (**b**) and minus 1 km (**c**) show similar agreement at different absolute values. Slight differences can be seen at background levels with increased agreement in 2010 for the increased lower integration limit (**b**). *This figure is added to the supplement of the revised version.*

the emission. We also clarify this in the revised version in the Methods section. We now provide a study in the supplement comparing the results for stratospheric SO_2 burden and sAOD, varying the lower integration limit by ± 1 km. We provide the figures here in the responses as Fig. 2 and 3, and additionally add them to the supplement. In the cited paragraph of the submitted manuscript, we state "... that total sAOD is highly sensitive to the altitude chosen as a lower cutoff". However, that does not mean that the agreement with observations is highly sensitive to the lower integration limit. We now clarify this in the manuscript, and refer to the supplement figure.

Line 556-561: I would make it clear in this paragraph (and for the whole discussion of AOD from Line 530 onward) that here you are talking about the sulfuric acid aerosol plume and the AOD resulting from these aerosols, whereas so far in the paper you talked about the SO_2 plume. The two SO_2 and sulfuric acid aerosol plumes likely look different.

We thank the reviewer for the suggestion, and made the transition from SO_2 to AOD and extinction clearer in the revised version, separating it into two different subsubsections, and defining the difference in the Methods.

Line 565/566: The first sentence of the discussion is not true. You do not show anything related to "aerosol formation". You only show comparison with AOD observations, but this does not tell you anything about aerosol formation processes.

The reviewer is correct, that we do not show aerosol formation here, and we removed it from the statement: "We showed that SO_2 emissions from explosive volcanic eruptions and the subsequent plume evolution can be reasonably reproduced in EMAC ..."

Line 575: You do not show "aerosol burden" here. Thus, you can not make any conclusions about this. Do you mean "forecasted" instead of "examined"? They can be examined, but just not immediately.

We reformulated to: "... and results in differing stratospheric lifetimes for the volcanic plume".
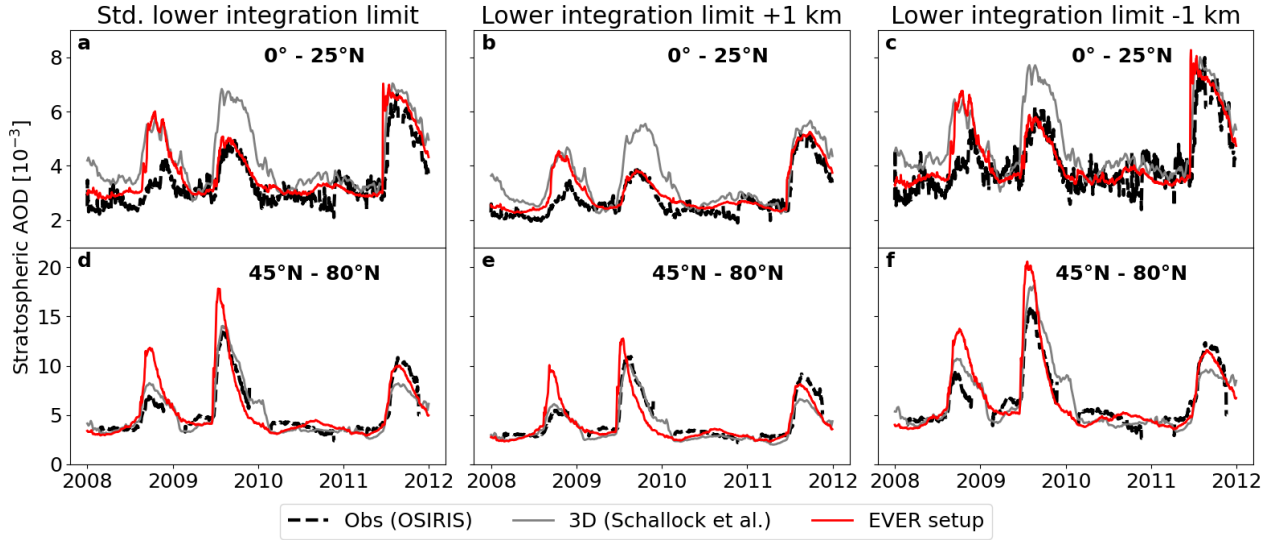
25

Figure 3: Sensitivity of the evaluation result to the lower integration limit (applied to observations and simulations) for the derivation of sAOD. **a,d** Lower integration limit as applied in the manuscript (16 km from 0-25°N and 12 km from 45-80°N). Variations of plus 1 km (**b,e**) and minus 1 km (**c,f**) show similar agreement at different absolute values. The most significant differences are observed for the Sarychev volcano, that shows strongly decreased simulated sAOD at increased lower integration limit at higher latitudes (**e**). *This figure is added to the supplement of the revised version.*

Regarding the second question: We refer to the 3D simulation here, that only injects 3D perturbations of stratospheric SO$_2$ a couple of days after the initial eruption. Thus, the short-term evolution cannot be examined. We clarified this in the revised version.

Line 589-593: You did not analyze how to "adequately simulate stratospheric aerosol burden". You cannot make any conclusions about stratospheric aerosol burden, if you do not show aerosol burden in the manuscript. The first sentence is confusing. What do you mean with "differences" in the first sentence of this paragraph? A difference compared to what? Again, what is the importance of chemical loss of SO$_2$ in the whole analysis? This could also be discussed here. You cannot make conclusions about the "sulfate" lifetime with the analysis shown in your manuscript.

We removed the term "aerosol burden" in the mentioned paragraph. We meant that the importance of each parameter is differing. We reformulated the sentence to: "The sensitivity studies revealed that the importance of accurately constraining the emission parameters for adequately simulating the volcanic SO$_2$ plume differs for each parameter."

Line 594-605: I agree that the horizontal extent of the emissions can influence the simulations. "...emissions are constrained to a single horizontal grid box in this study..." I know what you mean, but this reads wrong. You also applied column emissions and vertically gaussian distributed emissions, which do not inject into "one single grid box". I would change this to "...emissions are constrained to a single grid box or columns of gid boxes in this study..." or make this clearer in a different way (e.g. what is the difference between a horizontal and a vertical gid box?

) To me a grid box is a grid box... whether it is vertical or horizontal. "...leading to non-linearities in the model that diverge from reality..." This statement needs references. Why is this important? What non-linearities are you talking about? I recommend highlighting the impact on aerosol formation/microphysics from this artefact (e.g. Vattioni et al. 2024). In this paragraph you should also discuss the effect of the vertical and horizontal resolution of the model, since it is known that this can affect simulations of volcanic plumes. "This concentration can lead to lower SO_2 and aerosol mixing ratios in the mid- to long- term, as aerosols grow excessively large and subsequently sediment out of the stratosphere, as observed following the Nabro eruption." You provide an explanation for lower aerosol mixing ratios, but what would be the reason for differences in SO_2 mixing ratios? Also this sentence (and the whole paragraph) needs references, since you don't show this with your results.

We thank the reviewer for this remark, and apologize for being imprecise. We reformulated the paragraph according to the reviewer's suggestions, removed the part on the aerosols, and only concentrated on SO_2:

" However, column or point emissions come with inherent limitations as well. First, emissions are constrained to a single grid box or columns of grid boxes in this study, potentially resulting in localized and exaggerated mixing ratios of SO_2, when volcanic plumes span areas exceeding one grid box (Tilmes et al., 2023). This may lead to a depletion of the oxidants, and subsequently to a slower oxidation to H_2SO_4. While distributing emissions on multiple columns horizontally may mitigate this effect by defining multiple emission points per eruption (Tilmes et al., 2023), we did not explore this aspect in our analysis. However, in detailed studies of strong eruptions, we additionally recommend exploring the effects of emissions over multiple columns and an extended time period to avoid non-linearities due to very high concentrations (see also Sect. ...). "

Line 623-625: What about the importance of the stratospheric entry point?

We stated before in line 590: "Primarily, emitting an appropriate quantity of SO_2 at the correct altitude appears to be the most critical factor." We agree that we show, that the stratospheric entry point is also important in the evolution. However, especially from Fig. 7 of the submitted manuscript, it is obvious that the altitude and emitted mass have the largest impact. We extended the paragraph in the revised version as follows:

" This aligns with our conclusion that altitude and mass are the most crucial emission parameters, which were directly determined from the emission inventory and do not depend on the availability of IASI observations. However, we showed that the timing and geographical location of the stratospheric entry point can lead to additional uncertainties, that remain when no IASI observations are available. "

Line 634: "emission" is written twice.

We corrected that.

Line 691: Conclusions last paragraph: This paragraph should be put into future tense (and or conjunctive), since like it is written know one could think that this is already provided or underway.
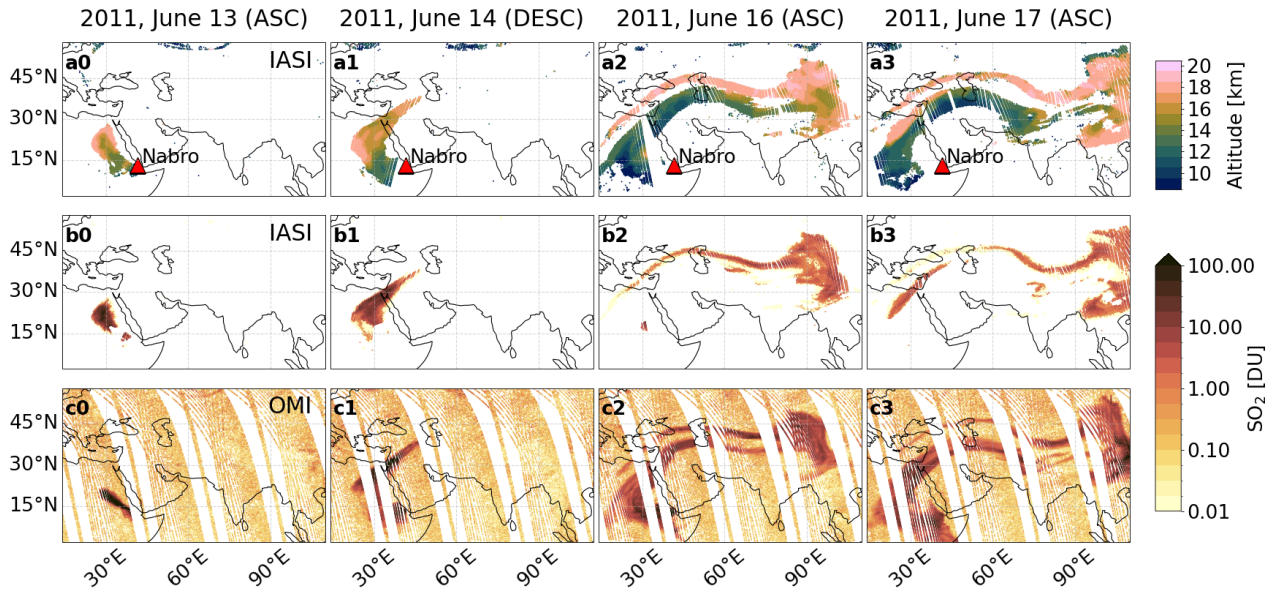
Figure 4: Observations of the Nabro plume on selected dates in the first week after the Nabro eruption. **a***, SO$_2$ plume height retrievals from IASI satellite observations (Clarisse et al., 2014) during ascending (ASC) and descending (DESC) orbits **b*,** Derived SO$_2$ columns from IASI (Clarisse et al., 2014) observations. **c*,** Derived SO$_2$ columns from OMI (Li et al., 2020) observations. IASI SO$_2$ columns are calculated assuming that all SO$_2$ is centered at the retrieved plume height (top row), and we only display pixels where the retrieved SO$_2$ plume is detected in the stratosphere (above 14 km), whereas OMI displays the total column. Note that the timing of the observations does not coincide in general.

Changed to "These may include ..."

Figure 1: It is not very helpful showing code in the main manuscript since this will not be helpful to most readers. If at all I would put this into a supplement.

We moved the namelist setup to the supplement.

Figure 2: From just looking at the figure caption it is not clear what the "plume" refers to: SO$_2$, Aerosol in general, ash or sulfuric acid aerosol? Maybe specify in the caption what the IASI satellite measures.

Thank you for the remark. We included this information in the caption. We additionally merged Figures 2 and 3 of the manuscript (see Fig. 4 here for the new version)

Figure 3: "amount" is not very specific. I would call the unit by its name (SO$_2$ column).

Replaced here and also in other occurrences.

Figure 4: The caption could be clearer. What do you mean by "shortly after"? If you compare observations to the altitude of the "maximum SO$_2$ mixing ratio", only one altitude should be displayed in your plot, since there is only one maximum in the vertical column, right? I am confused here. Maybe change the last sentence
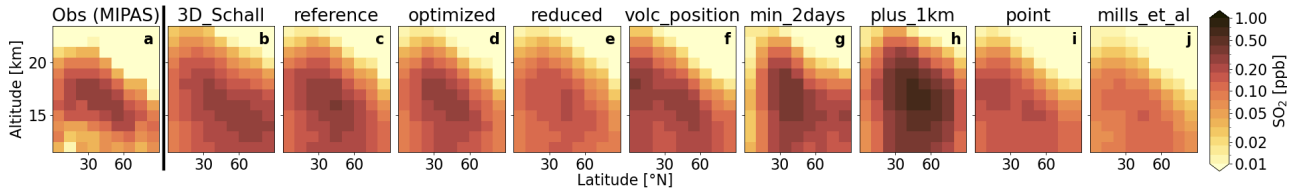
28

Figure 5: Zonally and 5-day averaged (July 14 - July 18, 2011) profile of northern hemispheric stratospheric SO$_2$ mixing ratios derived from the MIPAS satellite (**a**) is compared to the respective SO$_2$ mixing ratios simulated in the reference (**c**) and sensitivity simulations (**b,d-j**) for the same period (one month after the eruption of the Nabro volcano).

to: "In the simulations SO$_2$ was only injected into the stratosphere, except for mills_et_al".

We reformulated "shortly after the eruption (top) and three days later (bottom)" to "one day (top) and four days (bottom) after the initial eruption".
For each horizontal coordinate, there is one altitude referenced in the IASI dataset. From the model results, we also retrieve the altitude of maximum SO$_2$ per column. We reformulated to "... is compared with the altitude of maximum SO$_2$ mixing ratios for each vertical column from the sensitivity simulations ..." to clarify this.
For the last sentence, we used the suggestion from the reviewer.

Figure 5: See comments on Figure 4. Why don't you compare to OMI as well?

We adjusted the caption according to the previous comments. OMI includes observations from the troposphere and the stratosphere as it observes the total column, and the contributions cannot be distinguished. For this reason, we only compare to IASI, as we did not inject into the troposphere, and the focus of the study is on the stratosphere.

Figure 6: Maybe replace "zonally" with "zonally averaged". And also "study" with "simulations". I would skip "approximately" or be more specific. Does the date provided refer to the 5-day average or to the date of the eruption? There is no space between the first and the second panel, and the black line covers the "0". Please correct this.

We adapted the suggested adjustments, and clarified the time range. Additionally, we corrected the plot (see Fig. 5 here).

Figure 7: The first sentence of the caption can be skipped or integrated into the second one. What is the unit of the x axis? The format mm/dd is not used universal (in Europe dd/mm is more common). Thus, I suggest writing Jun 15, Jul 1, Jul 15 and so one, to make this clear. Y-Axis label: It is "SO$_2$", not "SO2". Why don't you show the spatial correlation for 15/6? Concerning the "stratospheric cutoff altitude": Do you mean tropopause? If not, why don't you use the tropopause altitude? If yes, I would name the tropopause by its name. Why did you choose these altitudes? Did you check the sensitivity of your assumptions? Looking at the satellite data and your simulations, you can see that 3 days after the eruption a considerable amount of the plume is exactly around 30°N. Thus, slightly changing the "stratospheric cutoff altitude" might have an impact on the results shown
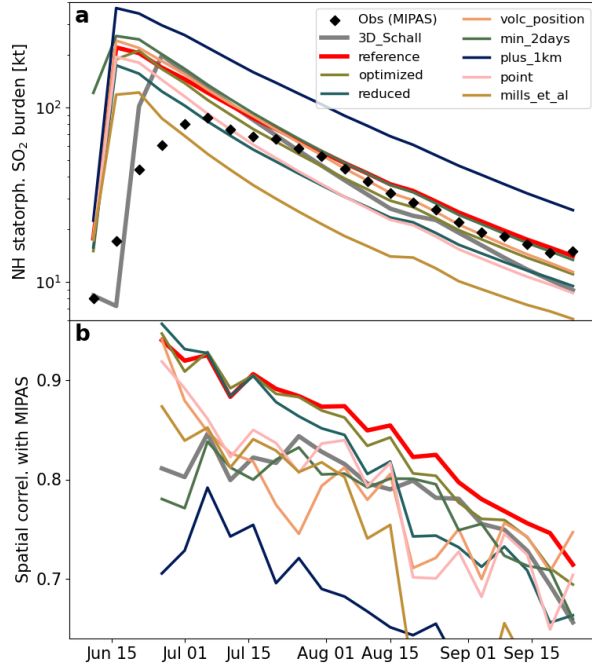
29

Figure 6: **a,** Total northern hemispheric stratospheric SO$_2$ burden as observed from the MIPAS satellite (black dots) and from the sensitivity simulations (5-day averages). **b,** Spatial correlation in the latitude-altitude plane between the simulations and MIPAS observations, i. e. the spatial correlation between the first and all other panels in Fig. 5.

We removed the first sentence, changed the format of the x axis label, and replaced "SO2" by "SO$_2$". We do not show the spatial correlation for the first two datapoints for two reasons. First, MIPAS only gives reliable data some time after the initial eruption, as described earlier in the manuscript (Line 355-359). Second, the 3D emissions are only applied ten days after the initial eruption. We could show the spatial correlation before, however, it would not tell us anything about the simulations. The updated figure can be found in Fig. 6.

We renamed "stratospheric cutoff altitude" to "lower integration limit" to avoid confusion in the revised manuscript. We did not use the tropopause here as the lower integration limit, because it is not necessarily coinciding in model and observations. To enable the comparison with observations, we used a universal lower integration limit for the calculation of the stratospheric SO$_2$ burden and the sAOD as defined here. This universal definition is now defined in the experimental setup section. To show that total sAOD and stratospheric SO$_2$ burden are sensitive to the lower integration limit, however the agreement is not, we now provide a study on the sensitivity to the lower integration limits in the supplement of the revised manuscript and here in Fig. 2 and 3.

30

We added the relative difference between the two simulations (see earlier reply and Fig. 1 here).

Figure 10 and 11: Same comment as on Figure 7. What does "stratospheric cutoff altitude mean" and how sensitive are results to this definition? Change to: "...using the EMAC model with the new *EVER* historic volcanic setup (red) and..." The axis label should read "SO2", not "SO$_2$"

We changed the caption according to the suggestions, and replaced "SO2" by "SO$_2$". We addressed the stratospheric cutoff altitude (lower integration limit) as outlined above. (see also Fig. 2 and 3)

Figures 11 and 12: To which wavelengths do the aerosol optical depths and extinctions refer to? This is crucial information which is missing.

We apologize for not providing the considered wavelength (750 nm), and provide it in the revised version.

Reply to comments of RC2:

The paper describes a new submodel within the MESSy framework for better simulation of mainly SO2 emissions from volcanic eruptions. Other potential applications are mentioned, such as emissions from wildfires, or volcanic degassing. In the study, the submodel is used to assess different distributions of SO2 emission distributions for two case studies. An explosive volcanic eruption (Nabro) and an effusive/surface emitting eruption (Kilauea). A setup for historical simulations from 2007 to 2011 is also presented. The model output is then discussed and compared with satellite retrievals.

The paper presents advances in modelling with the new submodel EVERv1.1 and is therefore fitting for GMD. EVER is a new tool to implement the emissions of SO2 from volcanic eruptions more flexibly and could be the basis for a more unified way to represent volcanic emissions in the future.

We thank the reviewer for the time to review our manuscript, the very constructive feedback, and overall positive evaluation. We report the comments (grey, bold) along with our replies (blue).

However, some restructuring of the text is needed to facilitate understanding. The manuscript is also very long and could be shortened considerably. I find the general structure of the presented manuscript very confusing. The methods used in this paper are generally stated but spread throughout the whole manuscript, which makes it difficult to read. I therefore recommend a more traditional structure with the introduction of the model, the different observations used in the analysis and the experiment description all in one dedicated chapter. Currently it reads more like the methods precede the respective result chapter and sometimes more information e.g. on the datasets is stated within the respective result sections. Also some information on the observational datasets is missing in the methodology. For example what gases are measured by what instrument. Some of this is explained later in the manuscript. There is also a section called "results" which would mean that everything before this is methodology? But this is not the case here. More detailed comments on this are below. This could also help making the paper more concise, since there are currently a few repetitions.

We applied the very helpful suggestions from the reviewer to our manuscript. First, we considerably shortened the manuscript. The main affected Sections are described below:

- The description of general impacts of volcanic eruptions were shortened in the introduction.

- We moved the namelist example for the new submodel (Fig. 1) and some of the detailed explanations to the supplement.

- We shortened the introduction of the Nabro and Kilauea volcano.

- We restructured the methods section, thereby avoiding repetitions.

- We summarized the satellite instruments in a structured table.

In addition, we generally restructured the methods and results section. Sect. 2, including the

presentation of the EVER submodel mostly remained in the same structure, only adding the description of the default namelist setup as an additional subsection. Sect. 3 now only focuses on the methods and the experimental setup, while Sect. 4 only shows the results of the three experiments. This results in the following new structure for Sect. 2-4:

Moreover, we provided the different observational instruments in a table, covering the most important aspects, and provide information which information is used in which results section in the experimental setup section 3.2 (see later specific comment).

The results show how different set ups of the emissions can influence the SO2 plume. It is discussed thoroughly, how the different SO2 injections influence the plume evolution after an eruption and finally is able to show, how the new set-up improves this with respect to observations. It is mentioned, that the Asian Monsoon influences the plume evolution, "simultaneously probing the dynamics in the model". This is not followed up with the necessary sensitivity simulations or discussion on the implications of dynamics and performance of the model.

We agree with the reviewer, that we did not address this topic. Although we initially planned to study this in detail, we concluded it to be out of scope of the manuscript, and did not include it. Sadly, the original sentence was kept in the submitted manuscript. We have removed it in the revised manuscript.

Since the code is not currently available, the results are currently not fully reproducible. However, the description in the paper should allow for a similar model to be constructed. But it is stated, that this submodel is portable to other base

models, while it was only tested for three very specific modal models. Would this submodel not function with a different type of sectional model?

Indeed, we can unfortunately not make the code public at the moment. We are working on changing this situation in the future (see also reply from Patrick Jöckel to the Chief Editor Comment).

We tested the submodel with the three mentioned modal aerosol models available in the MESSy environment. In general, there are no restrictions for the coupled aerosol model, but there is currently no sectional aerosol submodel available within the MESSy framework. For the coupling to other model systems, an interface has to be constructed. We clarified this in the manuscript.

Some more previously published research on the topic should be considered. E.g. different injection scenarios for the Pinatubo eruption in various models by Quaglia et al. (2023).

We added a more technical review on previous model studies on volcanic eruptions in the introduction, including more references to volcanic emission inventories (e. g. Carn et al., 2017; Diehl et al., 2012; Mills et al., 2016; Neely III and Schmidt, 2016; Schallock et al., 2023) and specific model studies (e. g. Brodowsky et al., 2021; Quaglia et al., 2023; Schallock et al., 2023; Vattioni et al., 2024).

While the title sufficiently describes the presented model, the abstract is currently still missing some comment on the models performance but rather only summarizes the methods. Some statement on the performance of the new model should be added here.

We have included information about the performance and the results of the sensitivity studies in the abstract of the revised version:

"This work documents the operation of a new submodel for tracer emissions from Explosive Volcanic ERuptions (EVER v1.1), developed within the Modular Earth Submodel System (MESSy, version 2.55.1). EVER calculates additional tendencies of gaseous and aerosol tracers based on emission source parameters, aligned to specific sequences of volcanic eruptions or other atmospheric emission sources, allowing for the employment of various vertical emission profiles. We show that volcanic SO_2 plumes can be reasonably reproduced through EVER emissions in numerical simulations with the ECHAM/MESSy Atmospheric Chemistry Model (EMAC), using satellite observations of SO_2 columns and mixing ratios following the explosive eruption of the Nabro volcano (Eritrea) in 2011 and a degassing event of the Kilauea volcano (2018) in Kilauea. Previous volcanic studies showed large variability in stratospheric SO_2 burdens depending on the chosen volcanic emission databases and parameters. Sensitivity studies on SO_2 emissions from the Nabro volcano explore perturbations of the emission source parameters, revealing that emission altitude and the emitted mass above the tropopause are most important for the mid- to long-term evolution of stratospheric SO_2 plumes, while the correct timing and geographical location of the stratospheric entrance is crucial for the short-term plume evolution. We integrate information from a volcanic SO_2 emission inventory, additional satellite observations, and our findings from the sensitivity studies to establish a historical standard setup for volcanic eruptions impacting stratospheric SO_2 from 1990 to 2023, successfully evaluated with satellite observations of stratospheric SO_2 burden and aerosol optical properties. We advocate for this to be a standardized setup in all simulations within the MESSy framework concentrat-

ing on the upper troposphere and stratosphere in this period. Further potential applications of EVER involve studies on volcanic ash, wildfires, solar radiation modification, and atmospheric transport processes."

Also please check all abbreviations, some of the abbreviations are not explained while others are specified several times. Similarly, some of the sources don't conform to the GMD format.

Thank you for this remark. We checked all abbreviations and sources.

Throughout the manuscript there are several very technical explanations and formulas, code snippets etc. that would fit better in the appendix or supplementary material. Figure 1 does not fit in this manuscript. This is only an example of a namelist, not a result and should be kept in the supplementary material. Particularly the first figure should represent the main results of the paper.

As this is a technical paper, we wanted to provide all information, that is necessary to work with the submodel. Example namelists are commonly presented in other GMD papers. However, as both reviewers suggest to remove the namelist from the manuscript, we moved it to the supplement in the revised version. Additionally, we moved the optimization algorithm for the Kilauea study to the supplement.

Line by line comments

L20ff/L55ff: Both of these paragraphs introduce heterogeneous chemistry on aerosol surfaces and ozone chemistry

We removed the second paragraph (L55ff), shortening the introduction.

"Up to 800 volcanic eruptions" why is there a limit? Is it impossible to simulate 801 eruptions?

This is just a technical limitation (as a maximum number has to be defined for memory usage), that can be adjusted. We agree, that this was not properly formulated, and we removed this statement.

L150: This formula is very general and not new to this paper, it should therefore be moved to the supplements.

We agree and removed the formula.

Fig1: The namelists are already in the supplementary material. This is also just an example and not a result. I suggest removing this figure and referring to supplementary material

We moved the namelist to the supplement (see general comments). We want to keep the example namelist in the supplement, as it is more concise than the complete historic namelist setup.

L143: Unexplained acronym AER

We removed this in the main text as it was specific to the namelist, that was now moved to the supplement.

L156: Why is this not possible?

One of the purposes of the TREXP submodel were artificial, mostly passive tracer experiments. For that reason, new tracer initialisation was possible in TREXP. In EVER we focus on volcanic species, that are chemically active and already treated in our chemical mechanism. Thus, we believed that it is not necessary for this submodel to include the definition of new tracers. As this information is not very important for the manuscript, however, we removed it.

L170ff: Is there a reference for how this is done in EVER 1.0?

Unfortunately, there is no reference for EVER 1.0. EVER 1.1 was developed while preparing the manuscript, as we implemented some improvements and adjustments for more flexibility. For that reason, we wanted to provide information on this change here.

L186ff: "and width" I find the term "width" misleading here as I understand that as spreading over several horizontal grid boxes. Or does this refer to the mass emitted? Vertical extent? If there is a fixed definition, please specify.

We apologize if the description was unclear. The term width referred to the sigma of the vertical Gaussian distribution. The mass is vertically distributed following a Gaussian distribution with the given sigma. It is only emitted in one horizontal grid box. We reformulated the whole paragraph to:

"In this distribution, the mass follows a Gaussian-shaped vertical profile with mean altitude and sigma defined in the namelist emission entry. It can be confined to a vertical extent as for the **Uniform distribution**, truncating the tail and scaling accordingly. The emission amount in each grid cell is calculated by considering the fraction of the error function integrated from the bottom to the top of the grid cell (for the lowermost and uppermost grid cells, the minimum and maximum altitudes are used, respectively) relative to the integral of the error function across the entire confined vertical extent."

L191ff: There are important microphysical implications for this, see e.g. in Fig. 3 in Tilmes et al. (2023)

We thank the reviewer for this suggestion. We included the reference (Tilmes et al., 2023) and adjusted the paragraph as follows:

"Although each emission point emits within a single grid box or a column of grid boxes, multiple emission points can be defined to accurately reproduce the horizontal distribution of a single eruption. This is especially important for large volcanic plumes entering the stratosphere, exceeding the area of one grid box (Schallock et al., 2023). Tilmes et al. (2023) showed, that the emission of the Pinatubo SO_2 plume over a horizontal region, rather than in one single column, leads to a significantly improved agreement with observations. However, in this study we only used one emission point for each eruption."

L201: What variables are nudged? What are the implications of nudging?

We clarify this in the revised version: "Temperature, the logarithm of the surface pressure, divergence and vorticity are "nudged" (more details in Jöckel et al., 2006) towards meteorological reanalysis data (ERA5, Hersbach et al., 2020) from the European Centre for Medium-Range weather forecasts (ECMWF). In addition, we employ the QBO submodel to weakly nudge the simulations to QBO zonal wind observations between 10 and 90 hPa (Giorgetta et al., 2002)

Table 1: Summary of satellite instruments and the respective observed quantities used for the evaluation of model simulations with the EVER submodel. More technical details on the observations and the applied retrievals can be found in the provided references.

Instrument	Observation method	Retrieved quant.	Opera-tion	Reference
Infrared Atmospheric Sounding Interferometer (IASI)	Thermal infrared emission from Earth (nadir)	SO_2 – columm + altitude retrieval	2006–now	Clarisse et al. (2014, 2012)
Ozone Monitoring Instrument (OMI)	UV & vis. solar backscatter radiation from Earth's surface (nadir)	SO_2 – total column	2004–now	Li et al. (2020)
Michelson Interferometer for Passive Atmospheric Sounding (MIPAS)	Mid-infrared emission spectrometer (limb)	SO_2 – 3D (stratosphere)	2002–2012	Höpfner et al. (2015, 2013)
TROPOspheric Monitoring Instrument (TROPOMI)	As OMI, extended to UV, visible, near-IR & shortwave	SO_2 – column (troposphere)	2017–now	Theys et al. (2017)
Optical, Spectroscopic and Infrared Remote Imaging System (OSIRIS)	Spectrally dispersed, scattered sunlight (limb)	**Aerosol extinction** – 3D (stratosph.)	2001–now	Rieger et al. (2019)

to avoid a phase drift."

L209: What does simplified chemistry mean, is OH prescribed?

The full mechanism for the simplified chemistry is provided in the supplement. The main difference for SO_2 is the direct conversion to H_2SO_4 on reaction with OH without intermediate products, such as SO_3. OH is not prescribed in any of the presented simulations. We discuss this in more detail in the revised manuscript:

"We use a simplified chemistry covering the basic tropospheric chemistry, including O_3, OH, NOx, NOy and basic sulfur chemistry (see supplement for details). Oxidation of SO_2 to H_2SO_4 is directly realized via reaction with OH, without producing any intermediates. We do not consider DMS and OCS here (see supplement), leading to a potential underestimation of background maritime SO_2 concentrations."

L223ff: The following paragraphs lack some consistency, consider also summarizing some of the key properties of these measurements in a table. E.g. resolution, extent, what gases are measured, time when they were/are in operation. Just from reading this, I am not sure where these measurements are used later. Is it about SO2 or aerosol, tropospheric or stratospheric?

Thank you for this suggestion. We summarised the key information in a table (see Table 1) in the revised manuscript, and give more information on which satellites are used for which purpose in the experimental setup.

L313: "width of 2km" again I find the use of the word "width" a bit confusing, do you mean vertical extent?

We reformulated to: "Column emission centred at altitude from the emission inventory minus 1 km (17 km); Gaussian vertical distribution with a sigma of 2 km, confined to the altitude range from 16 to 18 km; ..."

L360ff: This information on what limitations the different satellite instruments/ retrievals have fits better in the methods section with the description of the in-

struments/ datasets.

We agree. This part was moved to the experimental setup section (3.2) within the methods, and we only refer to it here.

L374: "AKM" was already defined previously, it is again defined several times throughout the manuscript.

We checked the manuscript for all occurrences, and made sure, that it is only defined once.

L376: "Vertical width"

We reformulated to: "The comparison between simulated and observed SO_2 distributions reveals some discrepancies, particularly in the vertical extent of the elevated SO_2 mixing ratios. After applying the AKM to the simulations, the SO_2 distributions exhibit a slightly larger vertical extent compared to the observations."

L473ff: This reads like it should be in the discussion/conclusion

We agree on that, and moved it (slightly adjusted) to the discussion.

L479ff: Parts of these section introduce more motivation for the study again, which belongs in the introduction. I also recommend combining the methodology with the ones in the other chapters.

We agree with the reviewer, and moved it to the new general methods section. We also generally restructured the methodology and the results (see reply to the general comments for the new structure).

L509: The title for this subsection is misleading, is everything before this methodology?

We summarized all results in one section, and separate them from the methods in the revised version to make the structure clearer (see reply to the general comments).

L595: What about availability of oxidants in these spatially confined plumes? There would also be some differences between simplified chemistry and a more sophisticated chemistry scheme

We agree, that there will be differences between the simplified and more sophisticated chemistry. However, in the discussed simulations, we only used the more sophisticated chemistry scheme. Regarding the availability of oxidants: If the plume is constrained to a smaller volume in the model, the volcanic SO_2 will also encounter relatively less OH overall, and the oxidation to H_2SO_4 will become slower. We shortly discuss this in the revised version:

"First, emissions are constrained to a single grid box or columns of grid boxes in this study, potentially resulting in localized and exaggerated mixing ratios of SO_2, when volcanic plumes span areas exceeding one grid box (Tilmes et al., 2023). This may lead to a depletion of the oxidants, and subsequently to a slower oxidation to H_2SO_4."

L634: "emission emission"

Fixed this repetition.

L639: How are eruptions after 2012 currently simulated?

After 2010 only climatological volcanoes of Diehl et al. (2012) are used for tropospheric degassing.

L915-925: Check sources/format

Thank you for the remark. We checked the format of the sources.

References

Brodowsky, C., T. Sukhodolov, A. Feinberg, M. Höpfner, T. Peter, A. Stenke, and E. Rozanov (2021). "Modeling the Sulfate Aerosol Evolution After Recent Moderate Volcanic Activity, 2008–2012". *Journal of Geophysical Research: Atmospheres* 126.23, e2021JD035472. DOI: `https://doi.org/10.1029/2021JD035472`.

Brühl, C., J. Lelieveld, P. J. Crutzen, and H. Tost (2012). "The role of carbonyl sulphide as a source of stratospheric sulphate aerosol and its impact on climate". *Atmospheric Chemistry and Physics* 12.3, pp. 1239–1253. DOI: `10.5194/acp-12-1239-2012`.

Brühl, C., J. Schallock, K. Klingmüller, C. Robert, C. Bingen, L. Clarisse, A. Heckel, P. North, and L. Rieger (2018). "Stratospheric aerosol radiative forcing simulated by the chemistry climate model EMAC using Aerosol CCI satellite data". *Atmospheric Chemistry and Physics* 18.17, pp. 12845–12857. DOI: `10.5194/acp-18-12845-2018`.

Carn, S., V. Fioletov, C. McLinden, C. Li, and N. Krotkov (2017). "A decade of global volcanic SO2 emissions measured from space". *Scientific reports* 7.1, p. 44095. DOI: `10.1038/srep44095`.

Clarisse, L., P.-F. Coheur, N. Theys, D. Hurtmans, and C. Clerbaux (2014). "The 2011 Nabro eruption, a SO$_2$ plume height analysis using IASI measurements". *Atmospheric Chemistry and Physics* 14.6, pp. 3095–3111. DOI: `10.5194/acp-14-3095-2014`.

Clarisse, L., D. Hurtmans, C. Clerbaux, J. Hadji-Lazaro, Y. Ngadi, and P.-F. Coheur (2012). "Retrieval of sulphur dioxide from the infrared atmospheric sounding interferometer (IASI)". *Atmospheric Measurement Techniques* 5.3, pp. 581–594. DOI: `10.5194/amt-5-581-2012`.

Diehl, T., A. Heil, M. Chin, X. Pan, D. Streets, M. Schultz, and S. Kinne (2012). "Anthropogenic, biomass burning, and volcanic emissions of black carbon, organic carbon, and SO$_2$ from 1980 to 2010 for hindcast model experiments". *Atmospheric Chemistry and Physics Discussions* 12, pp. 24895–24954. DOI: `10.5194/acpd-12-24895-2012`.

Giorgetta, M. A., E. Manzini, and E. Roeckner (2002). "Forcing of the quasi-biennial oscillation from a broad spectrum of atmospheric waves". *Geophysical Research Letters* 29.8, pp. 86-1-86–4. DOI: `https://doi.org/10.1029/2002GL014756`.

Günther, A., M. Höpfner, B.-M. Sinnhuber, S. Griessbach, T. Deshler, T. von Clarmann, and G. Stiller (2018). "MIPAS observations of volcanic sulfate aerosol and sulfur dioxide in the stratosphere". *Atmospheric Chemistry and Physics* 18.2, pp. 1217–1239. DOI: `10.5194/acp-18-1217-2018`.

Hersbach, H., B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita, G. De Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger, S. Healy, R. J. Hogan, E. Hólm, M. Janisková, S. Keeley, P. Laloyaux, P. Lopez, C. Lupu, G. Radnoti, P. de Rosnay, I. Rozum, F. Vamborg, S. Villaume, and J.-N. Thépaut (2020). "The ERA5 global reanalysis". *Quarterly Journal of the Royal Meteorological Society* 146.730, pp. 1999–2049. DOI: `https://doi.org/10.1002/qj.3803`.

Hommel, R., C. Timmreck, M. A. Giorgetta, and H. F. Graf (2015). "Quasi-biennial oscillation of the tropical stratospheric aerosol layer". *Atmospheric Chemistry and Physics* 15.10, pp. 5557–5584. DOI: `10.5194/acp-15-5557-2015`.

Höpfner, M., C. D. Boone, B. Funke, N. Glatthor, U. Grabowski, A. Günther, S. Kellmann, M. Kiefer, A. Linden, S. Lossow, H. C. Pumphrey, W. G. Read, A. Roiger, G. Stiller, H. Schlager, T. von Clarmann, and K. Wissmüller (2015). "Sulfur dioxide (SO$_2$) from MIPAS

in the upper troposphere and lower stratosphere 2002–2012". *Atmospheric Chemistry and Physics* 15.12, pp. 7017–7037. DOI: 10.5194/acp-15-7017-2015.

Höpfner, M., N. Glatthor, U. Grabowski, S. Kellmann, M. Kiefer, A. Linden, J. Orphal, G. Stiller, T. von Clarmann, B. Funke, and C. D. Boone (2013). "Sulfur dioxide (SO_2) as observed by MIPAS/Envisat: temporal development and spatial distribution at 15–45 km altitude". *Atmospheric Chemistry and Physics* 13.20, pp. 10405–10423. DOI: 10.5194/acp-13-10405-2013.

Jöckel, P., A. Kerkweg, A. Pozzer, R. Sander, H. Tost, H. Riede, A. Baumgaertner, S. Gromov, and B. Kern (2010). "Development cycle 2 of the Modular Earth Submodel System (MESSy2)". *Geoscientific Model Development* 3.2, pp. 717–752. DOI: 10.5194/gmd-3-717-2010.

Jöckel, P., H. Tost, A. Pozzer, C. Brühl, J. Buchholz, L. Ganzeveld, P. Hoor, A. Kerkweg, M. G. Lawrence, R. Sander, B. Steil, G. Stiller, M. Tanarhte, D. Taraborrelli, J. van Aardenne, and J. Lelieveld (2006). "The atmospheric chemistry general circulation model ECHAM5/MESSy1: consistent simulation of ozone from the surface to the mesosphere". *Atmospheric Chemistry and Physics* 6.12, pp. 5067–5104. DOI: 10.5194/acp-6-5067-2006.

Jost, A. (2021). "Determining the SO_2 emission rates of the Kilauea volcano in Hawaii using S5P-TROPOMI satellite measurements". Bachelor's Thesis. Johannes Gutenberg Universität Mainz.

Klobas, J. E., D. M. Wilmouth, D. K. Weisenstein, J. G. Anderson, and R. J. Salawitch (2017). "Ozone depletion following future volcanic eruptions". *Geophysical Research Letters* 44.14, pp. 7490–7499. DOI: 10.1002/2017GL073972.

Kremser, S., L. W. Thomason, M. von Hobe, M. Hermann, T. Deshler, C. Timmreck, M. Toohey, A. Stenke, J. P. Schwarz, R. Weigel, S. Fueglistaler, F. J. Prata, J.-P. Vernier, H. Schlager, J. E. Barnes, J.-C. Antuña-Marrero, D. Fairlie, M. Palm, E. Mahieu, J. Notholt, M. Rex, C. Bingen, F. Vanhellemont, A. Bourassa, J. M. C. Plane, D. Klocke, S. A. Carn, L. Clarisse, T. Trickl, R. Neely, A. D. James, L. Rieger, J. C. Wilson, and B. Meland (2016). "Stratospheric aerosol—Observations, processes, and impact on climate". *Reviews of Geophysics* 54.2, pp. 278–335. DOI: 10.1002/2015RG000511.

Li, C., N. A. Krotkov, P. Leonard, and J. Joiner (2020). *OMI/Aura Sulphur Dioxide (SO2) Total Column 1-orbit L2 Swath 13x24 km V003, Goddard Earth Sciences Data and Information Services Center (GES DISC) [data set]*. DOI: 10.5067/Aura/OMI/DATA2022.

Mills, M. J., A. Schmidt, R. Easter, S. Solomon, D. E. Kinnison, S. J. Ghan, R. R. Neely III, D. R. Marsh, A. Conley, C. G. Bardeen, and A. Gettelman (2016). "Global volcanic aerosol properties derived from emissions, 1990–2014, using CESM1(WACCM)". *Journal of Geophysical Research: Atmospheres* 121.5, pp. 2332–2348. DOI: 10.1002/2015JD024290.

Neely III, R. and A. Schmidt (2016). *VolcanEESM: Global volcanic sulphur dioxide (SO2) emissions database from 1850 to present - Version 1.0 [Data set]*. DOI: 10.5285/76ebdc0b-0eed-4f70-b89e-55e606bcd568.

Pöschl, U., R. von Kuhlmann, N. Poisson, and P. J. Crutzen (2000). "Development and intercomparison of condensed isoprene oxidation mechanisms for global atmospheric modeling". *Journal of Atmospheric Chemistry* 37.1, pp. 29–52. DOI: 10.1023/A:1006391009798.

Pringle, K. J., H. Tost, S. Message, B. Steil, D. Giannadaki, A. Nenes, C. Fountoukis, P. Stier, E. Vignati, and J. Lelieveld (2010). "Description and evaluation of GMXe: a new aerosol submodel for global simulations (v1)". *Geoscientific Model Development* 3.2, pp. 391–412. DOI: 10.5194/gmd-3-391-2010.

Quaglia, I., C. Timmreck, U. Niemeier, D. Visioni, G. Pitari, C. Brodowsky, C. Brühl, S. S. Dhomse, H. Franke, A. Laakso, G. W. Mann, E. Rozanov, and T. Sukhodolov (2023). "Interactive stratospheric aerosol models' response to different amounts and altitudes of SO_2 injection during the 1991 Pinatubo eruption". *Atmospheric Chemistry and Physics* 23.2, pp. 921–948. DOI: 10.5194/acp-23-921-2023.

Rieger, L. A., D. J. Zawada, A. E. Bourassa, and D. A. Degenstein (2019). "A Multiwavelength Retrieval Approach for Improved OSIRIS Aerosol Extinction Retrievals". *Journal of Geophysical Research: Atmospheres* 124.13, pp. 7286–7307. DOI: 10.1029/2018JD029897.

Ruder, S. (2016). "An overview of gradient descent optimization algorithms". *arXiv preprint arXiv:1609.04747*.

Sander, R., A. Baumgaertner, D. Cabrera-Perez, F. Frank, S. Gromov, J.-U. Grooß, H. Harder, V. Huijnen, P. Jöckel, V. A. Karydis, K. E. Niemeyer, A. Pozzer, H. Riede, M. G. Schultz, D. Taraborrelli, and S. Tauer (2019). "The community atmospheric chemistry box model CAABA/MECCA-4.0". *Geoscientific Model Development* 12.4, pp. 1365–1385. DOI: 10.5194/gmd-12-1365-2019.

Schallock, J., C. Brühl, C. Bingen, M. Höpfner, L. Rieger, and J. Lelieveld (2023). "Reconstructing volcanic radiative forcing since 1990, using a comprehensive emission inventory and spatially resolved sulfur injections from satellite data in a chemistry-climate model". *Atmospheric Chemistry and Physics* 23.2, pp. 1169–1207. DOI: 10.5194/acp-23-1169-2023.

Schmidt, A., M. J. Mills, S. Ghan, J. M. Gregory, R. P. Allan, T. Andrews, C. G. Bardeen, A. Conley, P. M. Forster, A. Gettelman, R. W. Portmann, S. Solomon, and O. B. Toon (2018). "Volcanic Radiative Forcing From 1979 to 2015". *Journal of Geophysical Research: Atmospheres* 123.22, pp. 12491–12508. DOI: 10.1029/2018JD028776.

Textor, C., H.-F. Graf, C. Timmreck, and A. Robock (2004). "Emissions from volcanoes". *Emissions of Atmospheric Trace Compounds*. Ed. by C. Granier, P. Artaxo, and C. E. Reeves. Dordrecht: Springer Netherlands, pp. 269–303. ISBN: 978-1-4020-2167-1. DOI: 10.1007/978-1-4020-2167-1_7.

Theys, N., I. De Smedt, H. Yu, T. Danckaert, J. van Gent, C. Hörmann, T. Wagner, P. Hedelt, H. Bauer, F. Romahn, M. Pedergnana, D. Loyola, and M. Van Roozendael (2017). "Sulfur dioxide retrievals from TROPOMI onboard Sentinel-5 Precursor: algorithm theoretical basis". *Atmospheric Measurement Techniques* 10.1, pp. 119–153. DOI: 10.5194/amt-10-119-2017.

Tie, X. and G. Brasseur (1995). "The response of stratospheric ozone to volcanic eruptions: Sensitivity to atmospheric chlorine loading". *Geophysical Research Letters* 22.22, pp. 3035–3038. DOI: 10.1029/95GL03057.

Tilmes, S., M. J. Mills, Y. Zhu, C. G. Bardeen, F. Vitt, P. Yu, D. Fillmore, X. Liu, B. Toon, and T. Deshler (2023). "Description and performance of a sectional aerosol microphysical model in the Community Earth System Model (CESM2)". *Geoscientific Model Development* 16.21, pp. 6087–6125. DOI: 10.5194/gmd-16-6087-2023.

Timmreck, C., G. W. Mann, V. Aquila, R. Hommel, L. A. Lee, A. Schmidt, C. Brühl, S. Carn, M. Chin, S. S. Dhomse, T. Diehl, J. M. English, M. J. Mills, R. Neely, J. Sheng, M. Toohey, and D. Weisenstein (2018). "The Interactive Stratospheric Aerosol Model Intercomparison Project (ISA-MIP): motivation and experimental design". *Geoscientific Model Development* 11.7, pp. 2581–2608. DOI: 10.5194/gmd-11-2581-2018.

Vattioni, S., A. Stenke, B. Luo, G. Chiodo, T. Sukhodolov, E. Wunderlin, and T. Peter (2024). "Importance of microphysical settings for climate forcing by stratospheric SO_2 injections as

modeled by SOCOL-AERv2". *Geoscientific Model Development* 17.10, pp. 4181–4197. DOI: 10.5194/gmd-17-4181-2024.